

Genetic Sequences by Electron Microscopy

Thesis by  
Louise Tsi Chow

In Partial Fulfillment of the Requirements  
For the Degree of  
Doctor of Philosophy

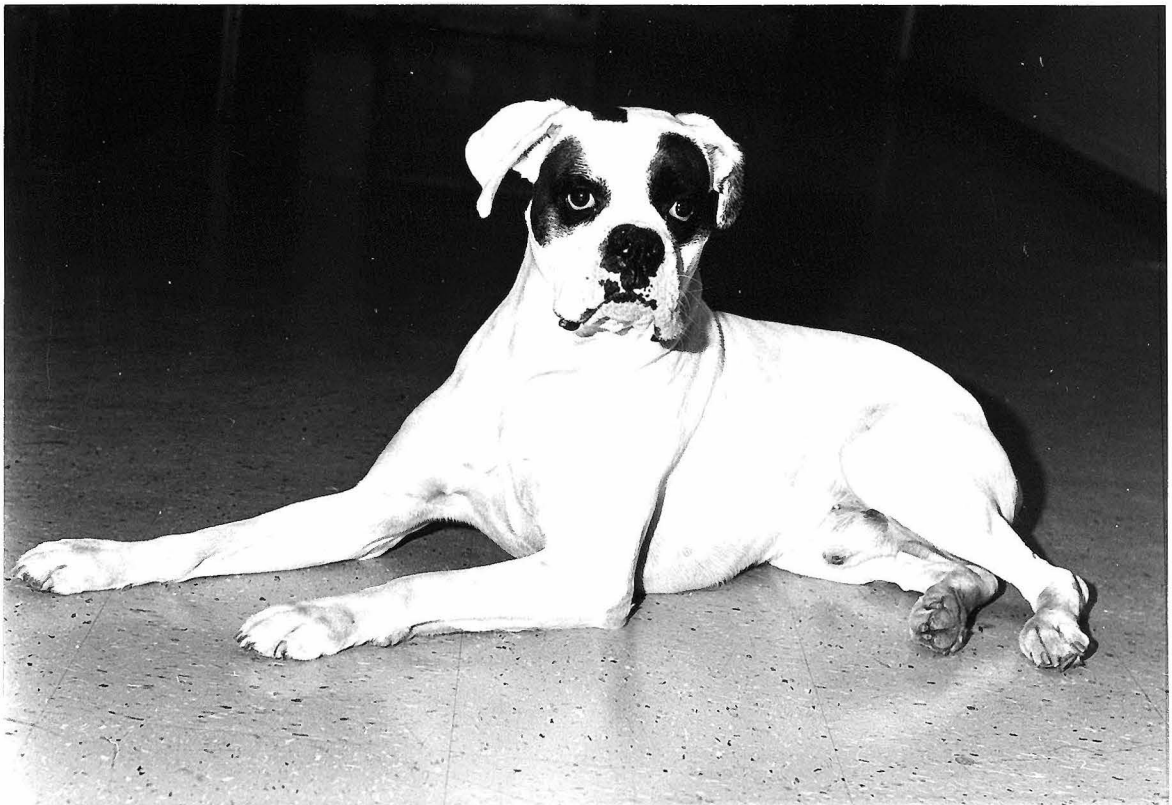
California Institute of Technology  
Pasadena, California

1973

( Submitted December 15, 1972 )

To My Parents





Embden Meyerhof

## Acknowledgments

I am very happy to have had the opportunity to be associated with many wonderful people throughout the years in Professor Norman Davidson's laboratory. Among them:

To Professor Norman Davidson, I owe my utmost gratitude not only for his continuous guidance, but also for his kindness, understanding, and encouragement, especially in trying times.

I deeply appreciate many fruitful discussions with Dr. Douglas E. Berg who introduces me to the "wonderful" world of  $\lambda$ dv's.

Thanks also go to Dr. LuBelle Boice who suggested the idea for the first part of this thesis.

I am indebted to Dr. Thomas R. Broker for his criticism and suggestions during the preparation of this thesis.

Last, but not least, I am grateful for the encouragement from members of my family.

## Abstracts

Part I      Sequence homology between the DNA molecules of the two temperate Bacillus subtilis bacteriophages, SPO2 and  $\phi$  105, has been mapped by electron microscope observation of heteroduplexes. There is a region of partial homology covering about 14 % of the genome (  $3.8 \times 10^6$  daltons ) close to the center of the heteroduplex molecule, flanked by completely non-homologous regions of lengths about  $1.2 \times 10^7$  and  $1.0 \times 10^7$  daltons on the two sides. Within the central homologous region there is a characteristic pattern of duplex regions and single-strand loops. The amount of duplex decreases as the denaturing power of the solvent used for preparing the electron microscope grids increases; this indicates that the DNA molecules of the two phages are only partially homologous within the homology region. The heteroduplex patterns show that there are no completely homologous nor completely non-homologous gene size sequences within the central region of partial homology. Since the phages are serologically related, we conclude that the antigenic determinants for serological cross-reactivity in the phage tails are coded for by genes in the central region of homology. This conclusion is consistent with available genetic data. Comparison of genetic and physical data indicates that the genes for DNA synthesis and for clear plaque formation in the two phages are non-homologous. The molecular weights of SPO2 DNA and  $\phi$  105 DNA are both calculated as  $26.3 ( \pm 0.3 ) \times 10^6$  daltons, from length measurements relative to  $\phi$  x 174 RF II DNA. Both DNA's have cohesive ends and are capable of reversible cyclization; the joined ends dissociate more readily than do those of  $\lambda$  DNA.

Our physical studies show that each phage DNA consists of a unique linear sequence and is not circularly permuted, in agreement with the conclusion from genetic studies that both phage maps are linear.

Part II      Circular duplex structures of the correct length are observed in the electron microscope in hybridization mixture of lysogen DNA and mature phage DNA for the case of the temperate Bacillus subtilis SPO2. This result shows that the sequence order of the prophage is a circular permutation of that of the mature phage. By making heteroduplexes of prophage DNA with that of the SPO2 deletion mutants R90 and S25, the att site of the phage has been mapped at  $61.2 \pm 0.6 \%$  from one end of the mature phage DNA, which has a length of 38,600 base-pairs. In the same coordinate system, the R90 deletion extends from  $58.9 \pm 0.7$  to  $66.8 \pm 0.8 \%$  on the SPO2 chromosome whereas the S25 deletion extends from  $63.2 \pm 0.6$  to  $66.9 \pm 0.7 \%$ . In similar experiments with lysogen and mature phage DNA's of the temperate B. subtilis phage,  $\phi 105$ , no circular structures were seen. This result shows that the sequence order in the prophage and the phage are collinear, without circular permutation. Short duplex segments, of length  $4830 \pm 250$  base pairs, with two single-strand arms at each end are seen at a low frequency after denaturation and renaturation of B. subtilis DNA. Several lines of evidence support the hypothesis that these duplex segments are formed by out-of-register renaturation of the 16s+23s ribosomal RNA genes (rDNA) of B. subtilis. They are of the correct length. Their formation is inhibited if homologous but not if heterologous

ribosomal RNA is added to the hybridization mixture. The frequency of occurrence of the duplex structures is consistent with the rDNA hypothesis. Heteroduplex molecules are seen with two or three rDNA duplex segments separated by single-strand substitution loops with specific lengths for each of the two single-strand arms of any one loop. On the basis these structures, linkage groups containing 7 to 9 rDNA sets ( each set containing one 16s and one 23s rDNA gene ) separated by spacer DNA's are proposed. The evidence indicates that if 5s rDNA is present in the set it is located near one end to give a gene order 16s-23s-5s. All of the 16s rDNA genes are linked to 23s rDNA and vice versa with little or no spacer DNA between a 16s and a 23s sequence. The spacer DNA between 23s and 5s must also be short. The prophage SPO2 bacterial att site maps at a distance 6200 bases away from a 16s+23s rDNA set which is itself separated by a very short spacer ( less than 600 bases ) from a second rDNA set.

Part III     The genetic sequences of seventeen different  $\lambda$ dv's obtained Dr. D.E. Berg have been mapped on  $\lambda$  DNA by the electron microscope heteroduplex method. The physical results are in agreement with the genetic analyses of the contents of the  $\lambda$ dv's. Different  $\lambda$ dv's exist predominantly either as monomers, or as dimers, or as trimers in recA<sup>-</sup> carrier cells. Higher order oligomers are also found in all  $\lambda$  dv preparations, Most of the  $\lambda$ dv's originated from  $\lambda$  phage carrying the nin deletion are observed to contain inverted partial duplications ( and are called amphimers ).

That is, the  $\lambda$ dv's carry a duplication of all or some of the sequences in an inverted order on the same strand. There are four types of inverted repeat structures: (1) complete inversion and duplication, (2) partial inversion and duplication with a long ( 5.7 % of  $\lambda^+$  ) non-inverted, non-duplicated ( or unique ) sequence on the left end covering the immunity region ( oriented according to the  $\lambda$  map ), (3) partial inversion and duplication with a long ( 5.1 % ) and a short ( 2.2 % ) unique sequence on the left and the right ends respectively, (4) partial inversion and duplication with a very short (  $<0.7$  % ) unique sequence on the right end of the  $\lambda$ dv's near the nin deletion. When the closed circular molecules of these inverted  $\lambda$ dv's are nicked lightly, denatured by alkali, and then reneutralized, the complementary sequences on the same DNA strand "snap back" to form double-stranded linear molecules. The unique sequences, if any, appear as single-stranded loops at one or both ends when the above treated DNA preparations are mounted by the formamide technique. Lambda phage that does not carry the nin deletion does not produce  $\lambda$ dv's with inverted repeats. Denaturation of the open circular molecules of these non-inverted  $\lambda$ dv's generates, as expected, only single-stranded circular and single-stranded linear molecules when mounted by the formamide technique. The nin deletion is believed to have caused somehow the formation of the  $\lambda$ dv's containing inverted repeats. The left end points of many  $\lambda$ dv's map around 73 % on the  $\lambda$  map while the right boundaries of many  $\lambda$ dv's map in the vicinity of the nin deletion which extends from 83.8 to 89.6 % on  $\lambda^+$ , indicative of two regions of high recombi-

nation frequency.

## Table of Contents

	<u>Page</u>
Acknowledgments . . . . .	iv
Abstracts . . . . .	v
Part I Map of the Partial Sequence Homology Between DNA Molecules of <u>Bacillus subtilis</u> Bacterio- phages SPO2 and $\phi$ 105 . . . . .	1
Part II Electron Microscope Mapping of the Phage <u>att</u> Sites for Bacteriophages SPO2 and $\phi$ 105, and of the Distribution of Ribosomal Genes on the <u>Bacillus subtilis</u> Chromosome . . . . .	18
Part III Electron Microscope Mapping of $\lambda$ dv DNA's . . . . .	88
Propositions . . . . .	187



Part I

Map of the Partial Sequence Homology Between DNA  
Molecules of Bacillus subtilis Bacteriophages SP02 and  $\phi$ 105

Louise T. Chow, LuBelle Boice\* and Norman Davidson

Division of Chemistry and Chemical Engineering

California Institute of Technology

Pasadena, California, U.S.A. 91109

\*Department of Molecular Biology and Biochemistry

University of California at Irvine

Irvine, California, U.S.A. 92664

*J. Mol. Biol.* (1972) **68**, 391–400

## Map of the Partial Sequence Homology between DNA Molecules of *Bacillus subtilis* Bacteriophages SPO2 and $\phi$ 105

LOUISE T. CHOW, LUBELLE BOICE† AND NORMAN DAVIDSON

*Gates, Crellin, and Church Laboratories of Chemistry  
California Institute of Technology  
Pasadena, Calif. 91109, U.S.A.*

and

*†Department of Molecular Biology and Biochemistry  
University of California at Irvine  
Irvine, Calif. 92664, U.S.A.*

(Received 14 February 1972)

Sequence homology between the DNA molecules of the two temperate *Bacillus subtilis* bacteriophages, SPO2 and  $\phi$ 105, has been mapped by electron microscope observation of heteroduplexes. There is a region of partial homology covering about 14% of the genome ( $3.8 \times 10^6$  daltons) close to the center of the heteroduplex molecule, flanked by completely non-homologous regions of lengths about  $1.2 \times 10^7$  and  $1.0 \times 10^7$  daltons on the two sides. Within the central homologous region there is a characteristic pattern of duplex regions and single-strand loops. The amount of duplex decreases as the denaturing power of the solvent used for preparing the electron microscope grids increases; this indicates that the DNA molecules of the two phages are only partially homologous within the homology region. The heteroduplex patterns show that there are no completely homologous nor completely non-homologous gene size sequences within the central region of partial homology. Since the phages are serologically related, we conclude that the antigenic determinants for serological cross-reactivity in the phage tails are coded for by genes in the central region of homology. This conclusion is consistent with available genetic data. Comparison of genetic and physical data indicates that the genes for DNA synthesis and for clear plaque formation in the two phages are non-homologous. The molecular weights of SPO2 DNA and  $\phi$ 105 DNA are both calculated as  $26.3 (\pm 0.3) \times 10^6$  daltons, from length measurements relative to  $\phi$ X174 RFII DNA. Both DNA's have cohesive ends and are capable of reversible cyclization; the joined ends dissociate more readily than do those of  $\lambda$  DNA. Our physical studies show that each phage DNA consists of a unique linear sequence and is not circularly permuted, in agreement with the conclusion from genetic studies that both phage maps are linear.

### 1. Introduction

Classical criteria for defining "related bacteriophages" are discussed by Adams (1959). These criteria include serology, size and morphology, chemical composition, genetic complementation and genetic recombination. Relatedness between organisms can also be estimated by DNA sequence homology measurements by the several kinds of DNA-DNA hybridization experiments. In particular, the electron microscope heteroduplex method (Davis & Davidson, 1968; Westmoreland, Szybalski & Ris, 1969;

† University of California at Irvine, Calif. U.S.A.

Davis, Simon & Davidson, 1971; Davis & Hyman, 1971; Fianndt, Hradecna, Lozeron & Szybalski, 1971) may be used not only to estimate the amount of homology but to map the homologous regions along the genome.

The lambdoid phages of *Escherichia coli* form such a family (Jacob & Wollman, 1961); their relatedness is indicated by morphology, serology, and genetic recombination. The regions of homology in various lambdoid phages have been mapped by electron microscopy (Simon, Davis & Davidson, 1971; Fianndt *et al.*, 1971). The well-known relatedness of the T-even phages has also been studied by electron microscope mapping methods (Kim & Davidson, 1971 Abst. 15th Annual Meeting Biophys. Soc. p. 266a) as well as by classical hybridization techniques (Bolton *et al.*, 1963; Cowie & Avery, 1970; Schildkraut, Weirzchowski, Marmur, Green & Doty, 1962).

Bacteriophages T3 and T7 are related by all the criteria mentioned above except that, since they exclude each other on infection, there is very little genetic recombination. (Viable recombinants have however been constructed (Hausman, Magalhaes & Araujo, 1961).) Their sequence homology has been studied by electron microscopy (Davis & Hyman, 1971) and by other DNA-DNA hybridization methods (Schildkraut *et al.*, 1962). The *micrococcus luteus* phages N1 and N6 are related serologically and by genetic recombination and appear to be practically identical by electron microscope and membrane filter DNA-DNA hybridization experiments (Lee & Davidson, 1970; Naylor & Burgi, 1956; Scalletti & Naylor, 1959).

Relatedness of the temperate *B. subtilis* bacteriophages SPO2 and  $\phi$ 105 has been proposed on the basis of serology and morphology (Boice, Eiserling & Romig, 1969; Boice, 1969). We report here our electron microscope studies of the degree and location of the sequence homology between the DNA molecules of these two phages.

## 2. Experimental Procedure

### (a) Bacteriophages

SPO2 and  $\phi$ 105 were grown as described by Boice *et al.* (1969).

### (b) Electron microscopy

DNA heteroduplexes were prepared by treating a mixture of the two phages with alkali and EDTA, thus causing simultaneous lysis of the phage and denaturation of the released DNA's, followed by reneutralization with Tris-HCl and renaturation in a formamide solvent (Davis *et al.*, 1971). Details of the procedure are as follows. A few  $\mu$ l. of each phage solution in CsCl, sufficient to give 0.25  $\mu$ g of each kind of DNA, was added to a solution containing 5 to 15  $\mu$ l. of 1 M-NaOH, 5  $\mu$ l. of 0.2 M-trisodium EDTA (pH 8), and water so that the final volume was 35 to 45  $\mu$ l. After standing at room temperature for 15 min, this denatured DNA solution was neutralized by adding a volume of 0.2 M-Tris-OH, 1.8 M-Tris-HCl equal to the volume of NaOH solution previously added. The final pH is 8.5 and the final volume 50  $\mu$ l. This is followed by the addition of 50  $\mu$ l. of formamide. Renaturation was allowed to proceed for about 40 min at room temperature; the DNA solution was then dialyzed into 0.01 M-Tris, 0.001 M-trisodium EDTA (pH 8.0) at 4°C.

The spreading solution for electron microscopy contained final concentrations of 0.076 M-Tris-OH, 0.024 M-Tris-HCl (pH 8.6), 0.01 M-trisodium EDTA, 0.05 M-ammonium acetate, 50  $\mu$ g cytochrome c/ml., about 0.5  $\mu$ g DNA/ml., and 50, 40, or 30% (by vol.) of formamide, to a final volume of 0.100 ml. One-half of this is used per spreading. The hypophase contained 0.0066 M-Tris-OH, 0.0034 M-Tris-HCl, 0.001 M-trisodium EDTA, and 20, 10, and 5% formamide, respectively. The significance of the variation of the formamide concentration is discussed later. Self-renatured samples were prepared by the same procedure with just one phage.

Samples mounted by the "aqueous" technique for the investigation of circle formation

were spread from 0.5 M-ammonium acetate with cytochrome *c* at 50 to 100  $\mu$ g/ml. onto 0.25 M-ammonium acetate.

Single-strand and duplex DNA length measurements were standardized against single-strand  $\phi$ X174 DNA and double-strand  $\phi$ X174 RFII DNA mounted on the same grid, with molecular weights taken as  $1.72 \times 10^6$  and  $3.45 \times 10^6$  daltons, corresponding to 5200 nucleotides or nucleotide pairs, respectively (Davidson & Szybalski, 1971).

(c) *Strand separation*

The procedure described by Summers & Szybalski (1968) as modified by Davis & Hyman (1971) was used.

### 3. Results and Discussion

(a) *Molecular lengths and weights*

The measured molecular lengths and weights are shown in Table 1.

TABLE 1  
*Duplex molecular lengths† and weights‡*

	SPO2	$\phi$ 105
No. of molecules in sample	18	20
Average length in $\phi$ X174 RF units	7.65( $\pm 0.11$ )	7.59( $\pm 0.08$ )
Molecular wt‡ ( $\times 10^{-6}$ )	26.4( $\pm 0.4$ )	26.2( $\pm 0.3$ )

† These are molecular lengths of homoduplexes prepared by denaturation and renaturation and mounted from 40% formamide. Only completely double standard molecules were measured. About 20% of the strands in either phage preparation contained a break.

‡ Calculated from relative length measurements taking  $\phi$ X RFII as  $3.45 \times 10^6$  daltons (see text).

The important conclusion from the data in Table 1 is that, within the uncertainty of electron microscopy  $\phi$ 105 and SPO2 DNA's have equal molecular weights of  $26.3 (\pm 0.4) \times 10^6$  daltons.

Birdsell, Hathaway & Rutberg (1969) report  $24$  to  $26 \times 10^6$  daltons for  $\phi$ 105 DNA by band sedimentation and  $21$  to  $24 \times 10^6$  daltons by absolute contour length measurements. Boice *et al.* (1969) give  $26 \times 10^6$  daltons for SPO2 DNA by the latter method. The results are all in reasonably good agreement.

(b) *Cohesive ends*

Birdsell *et al.* (1969) report that in grids of native  $\phi$ 105 DNA prepared by spreading from aqueous ammonium acetate about 20% of the molecules were circular. They report seeing many molecules with ends in close opposition. This observation is characteristic of molecules that have cyclized by joining of cohesive ends; the spreading forces in the formation of the protein film often cause dissociation of the joined ends. These authors, however, report that annealing samples at  $45^\circ\text{C}$  in 0.6 M-NaCl did not result in end joining, although these conditions do lead to efficient joining of lambda ends. On the other hand, Romig (1968) reports that the DNA of SPO2 can be circularized by heating to  $75^\circ\text{C}$  in 0.6 M-NaCl and slow cooling. Boice *et al.* (1969)

observed "several" circular molecules and a linear dimer molecule in a sample containing many linear molecules in aqueous ammonium acetate spreads of SPO2 DNA.

Thus, the indications were that both these phage DNA's were capable of cyclization via cohesive ends. We observe that in grids spread by the aqueous technique after dialysis of either native or self-renatured DNA from 50% formamide as described in Experimental Procedure (thus including exposure to renaturing conditions), about 40 to 50% of SPO2 and of  $\phi$ 105 DNA molecules are circles. Some linear and circular dimers are seen also. In spreading from 30% formamide onto 5% formamide, pulled apart circles (i.e. molecules with ends close to each other) are seen at a high frequency. Only linear molecules with ends well separated are seen in spreads from 40% formamide onto 10% formamide or from 50% formamide onto 20% formamide. However, if one starts with hydrogen-bonded circular  $\lambda$  DNA molecules, spreading from 40% formamide onto 10% gives circular molecules or pulled apart circles predominantly. Examples of circular molecules and of pulled apart circles are shown in Plate I.

The observations show that  $\phi$ 105 and SPO2 DNA have cohesive ends, but that the joined ends are more readily dissociated than are joined  $\lambda$  ends. Thus, the cohesive ends of  $\phi$ 105 and SPO2 DNA are shorter than those of  $\lambda$  (i.e. 12 nucleotides) and/or they have a higher A · T base composition (the  $\lambda$  ends having 83.3% G + C; Wu & Taylor, 1971). It should be noted that we have no evidence as to whether or not the ends of  $\phi$ 105 and of SPO2 DNA's can cohere to each other.

Only linear duplexes are seen in spreadings of self-renatured SPO2 and  $\phi$ 105 DNA from 40 or 50% formamide. This is physical evidence that each of the DNA's consists of a unique linear sequence and is not circularly permuted.

#### (c) *Heteroduplexes of SPO2 and $\phi$ 105 DNA*

No molecules containing both duplex regions and the typical deletion loops, substitution loops or forks, that characterize heteroduplexes between molecules which are homologous in some regions and non-homologous in others, were seen in our first attempts to make heteroduplexes by lysing a mixture of equal amounts of the two phages, denaturing and renaturing. There were some single-strand nicks in each preparation. These broken strands gave renatured molecules with a duplex region joining a single single-strand region. Some perfect duplexes and some unrenatured single strands were seen.

Several interpretations seemed possible.

- (1) The DNA's are completely non-homologous.
- (2) The DNA's are sufficiently homologous to appear duplex throughout. This, for example, is the case for the *M. luteus* phages, N1 and N6 (Lee & Davidson, 1970).
- (3) There are regions of homology but they occupy such a small fraction of the genome that the rate of renaturation to form heteroduplexes is very small compared to the rate of renaturation to form the respective homoduplexes.

A positive way to settle the question is to separate the complementary strands of each phage DNA and to observe the products of renaturation of  $l$  strands of one phage DNA with  $r$  strands of the other. However, we could not achieve any strand separation by poly(U,G) binding; only one sharp band was obtained for either denatured SPO2 or  $\phi$ 105 DNA. (The same technique in our hands gave a good separation of the complementary strands of  $\lambda$  DNA.)

It was then discovered by very careful scrutiny of the grids prepared by denaturing

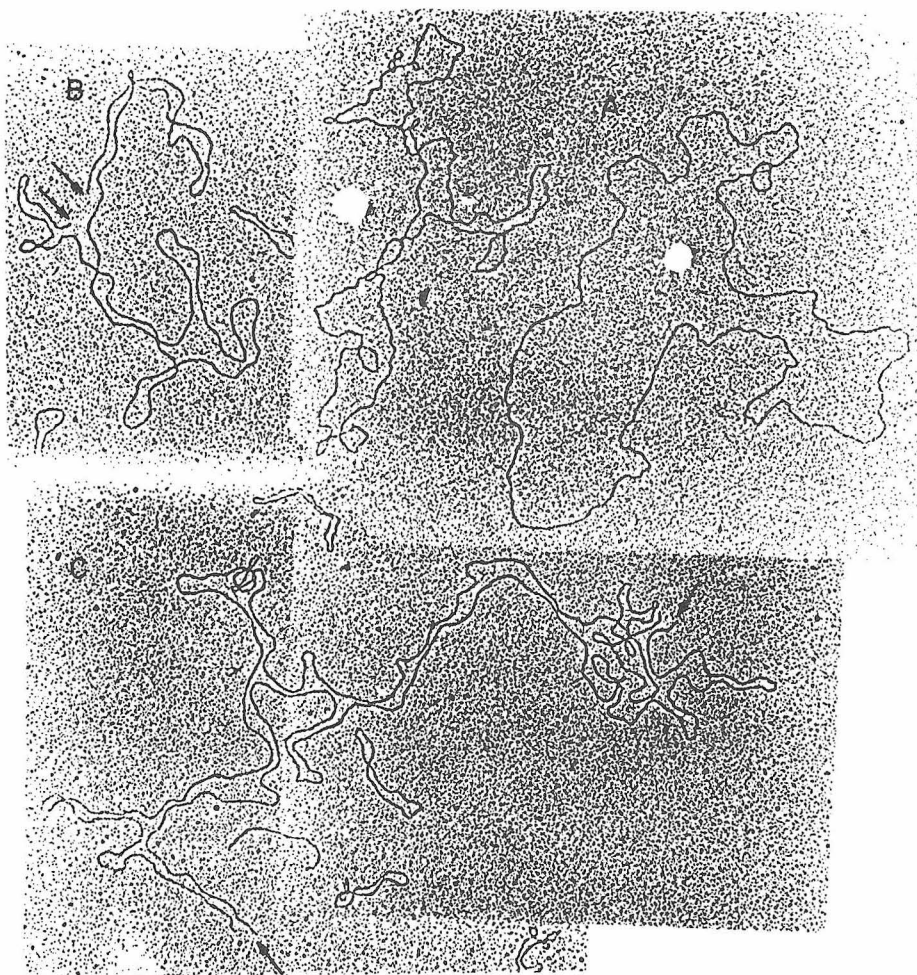


Plate I . Circular and pulled apart circular molecules .  
 (a) Circular SP02 molecules. The DNA was self-renatured and mounted from 0.5 M onto 0.25 M ammonium acetate as describe in the text. (b) and (c) Pulled apart circular monomer (b) and circular dimer (c) molecules. A mixture of phages SPO2 and  $\phi 105$  was denatured, renatured and then mounted from 30 onto 5% formamide as described in the text. The small circles are  $\phi x 174RFII$  added as length measurement standards. Arrows indicate the pulled apart ends .

Plate II. Heterocuplwx between SPO2 and  $\phi$ 105 DNA mounted from 40% formamide onto 10%. This is the tenth molecule from the top of Fig. 2 (b). A tracing of the center of the molecule is shown. Two of the non-homology loops in the center of the molecule are identified as *sa* and *sb*; the right and left non-homology regions are identified as *sr* and *sl* ( see Figs 1 and 2 (b) also ). The two *sl* strands and one of the *sr* strands are partially duplex due to renaturation with broken pieces of DNA of the same kind. As may be seen from the sizes of the  $\phi$ X 174 single strands and of the  $\phi$ XPF duplexes in the several pates, the magnification in Plate II is less than in Plates III and IV.



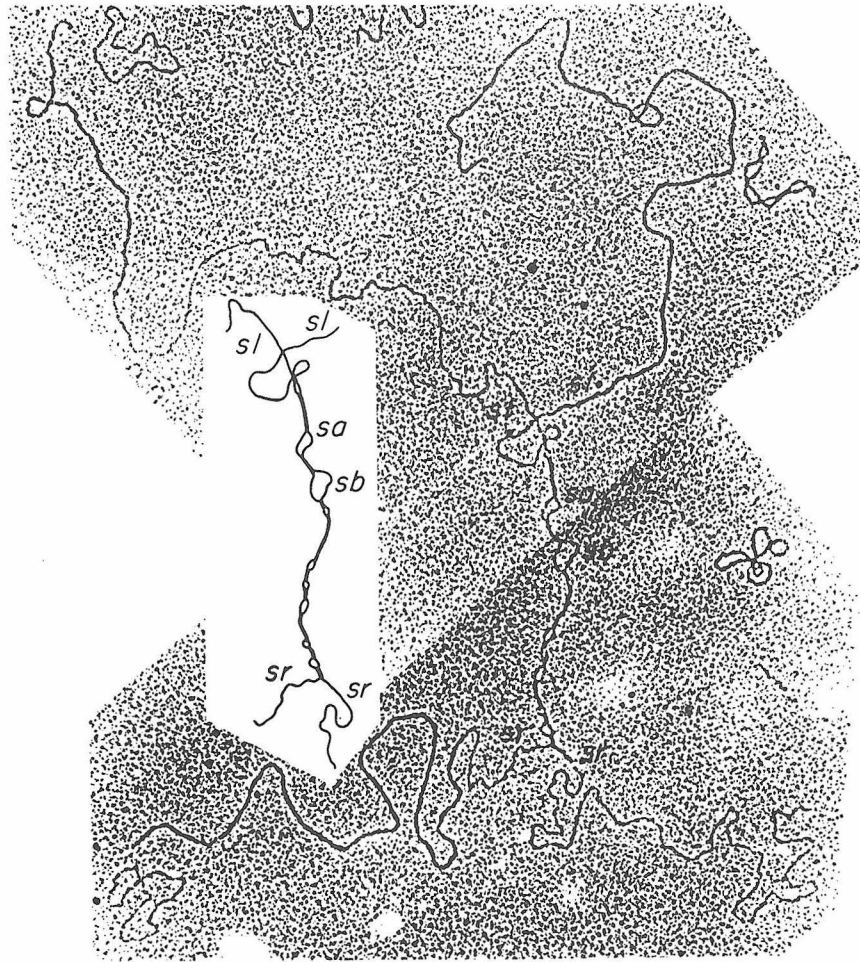


Plate II



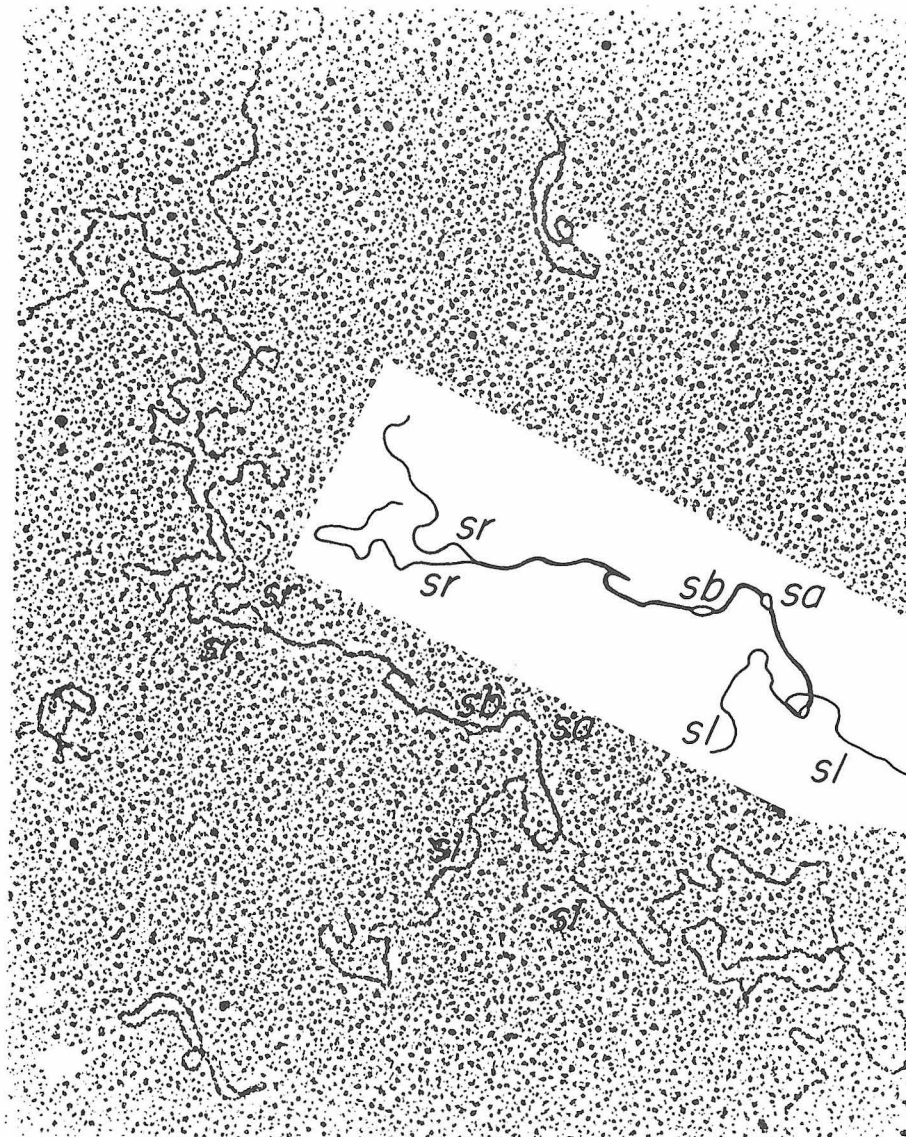


Plate III. SPO2/ $\phi$ 105 heteroduplex spread from 30% formamide onto 5%. This is the top molecule in Fig. 2(c). Symbols as in the legend to Plate II.

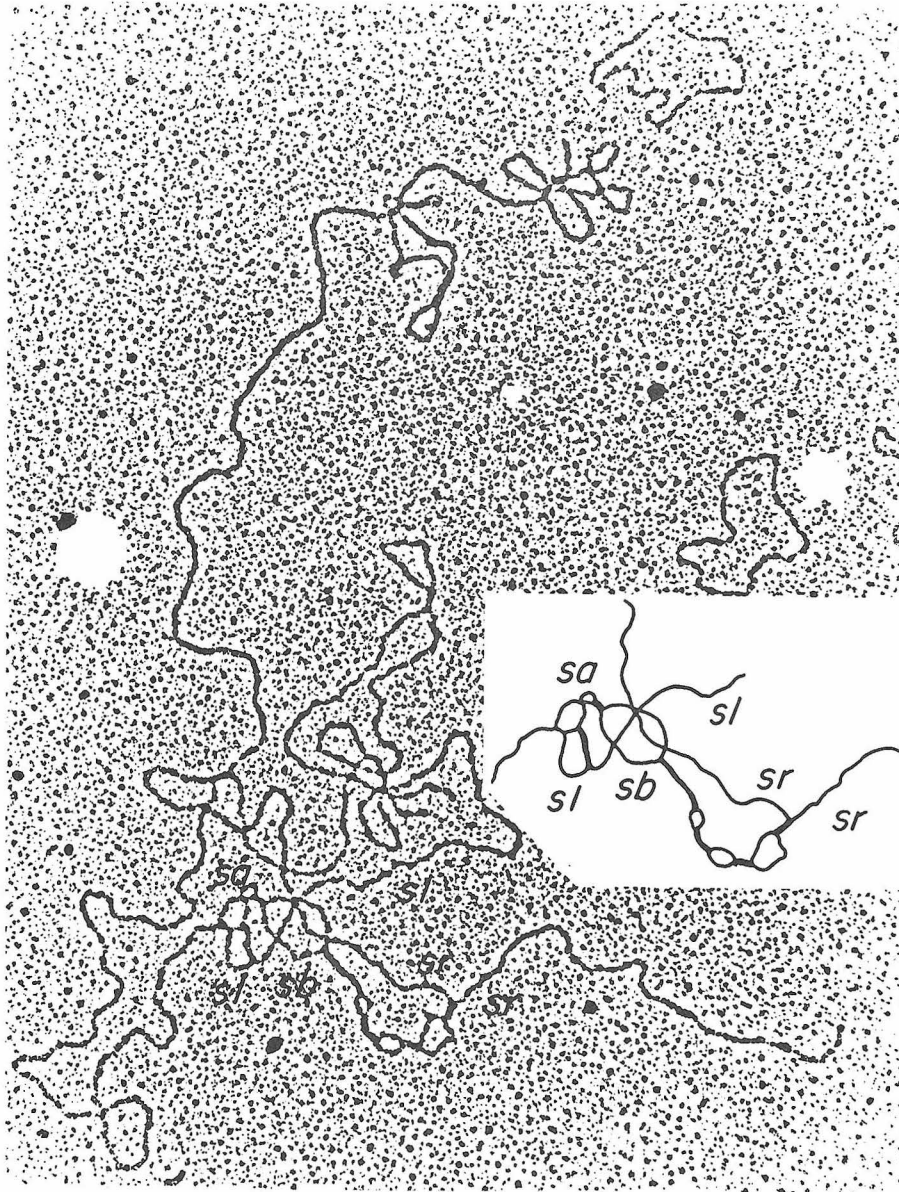


Plate IV. SPO2/ $\phi$ 105 heteroduplex spread from 50% formamide onto 20%. This is the eighth molecule from the top in Fig. 2 (a). Symbols as in the legend to Plate II. The top *sr* strand is partially duplex due to renaturation with a broken piece of DNA of the same kind.

and renaturing a mixture of the two kinds of DNA that heteroduplexes were present at a frequency of about 0.1 to 0.4% of the homoduplexes. The heteroduplexes contain a region of homology of length about  $3.8 \times 10^6$  daltons in the middle of the molecule, with two dangling single strands on each side, with lengths corresponding to duplex molecular weights of about  $1.2 \times 10^7$  and  $1.0 \times 10^7$  daltons, respectively. The homology region shows a pattern of duplex regions spaced by single-strand loops; several typical molecules are shown in Plates II, III and IV. An over-all summary of the structure is shown in Figure 1. No such heteroduplexes were seen in self-renatured preparations of either kind of DNA.

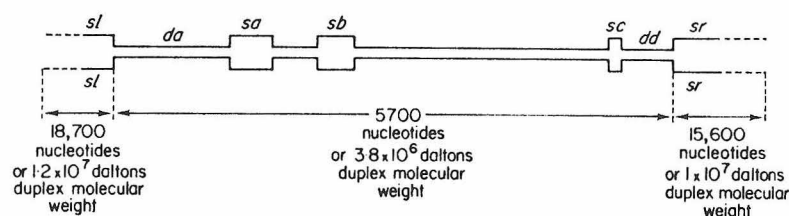


FIG. 1. Schematic representation of the heteroduplex structure between SPO2 and  $\phi$ 105 DNA's. The drawing depicts the pattern most frequently observed for heteroduplexes when spread from 30% formamide onto 5%; (see text and Fig. 2(c)). For purposes of identification in the micrographs and in the text, non-homology regions are labeled *sl*, *sa*, *sb*, *sr*, and duplex regions are labeled *da* and *dd*.

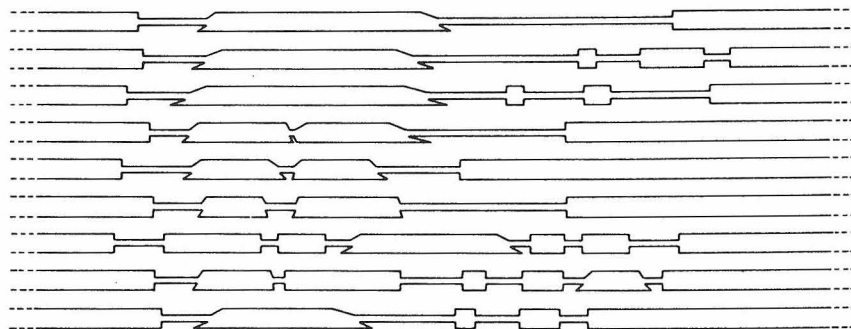
The region of non-homology on the left side of the molecule in Figure 1 is labeled *sl*. Some molecules were seen in which both single strands on this side had the same length of 18,700 ( $\pm 1000$ ) nucleotides ( $3.60 \phi X$ ). We therefore believe that the SPO2 and  $\phi$ 105 strands are of the same length on this side. In many cases, one single strand or both were shorter than this due to single-strand breaks. Similarly, the lengths of the non-homologous single strands on the right side, labeled *sr*, are both equal to 15,600 ( $\pm 800$ ) nucleotides.

The size and/or the number of single-strand loops within the central region of non-homology increases as the denaturing power of the solvent used for mounting the molecules in a basic protein film increases. Thus, Figure 2(c) shows the loop pattern for molecules in spreading from 30% formamide onto 5% formamide. In these heteroduplexes, the homology region appears as mainly duplex; two single-strand loops, *sa* and *sb*, on the left side of the homology region, each of length 100 to 400 nucleotides, are seen regularly; a smaller non-homology loop, *sc*, on the right side of the duplex region is seen frequently but not always.

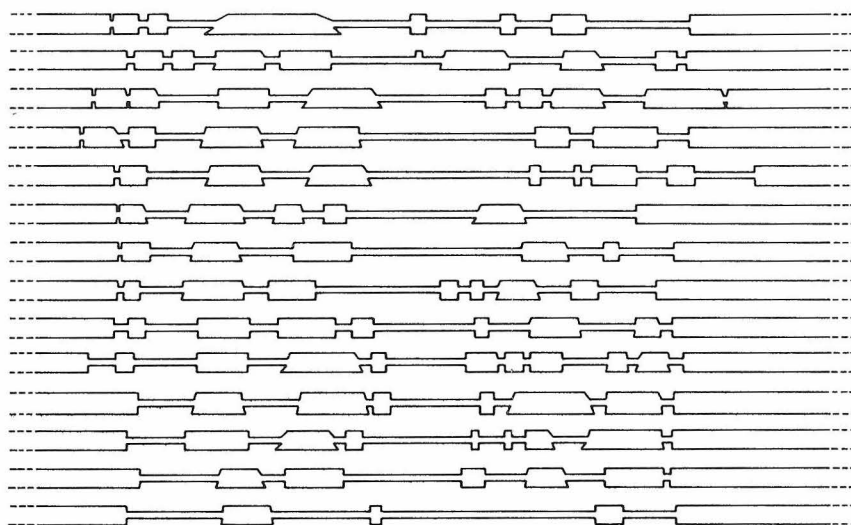
Figure 2(b) shows maps of molecules spread from 40% formamide onto 10%; Figure 2(a) is for 50 onto 20%. The loop patterns are not perfectly reproducible between different molecules, but there is a generally consistent and reproducible pattern. The estimated lengths of duplex regions and the total lengths between the left end of the first duplex segment and the right end of the last duplex segment under the three spreading conditions were: (a) 50% formamide onto 20%; 1500 to 2500 nucleotide pairs duplex, 3200 to 5000 nucleotide pairs total length; (b) 40% formamide onto 10%, 2100 to 3900 nucleotide pairs duplex, 4500 to 5500 total length; and (c) 30% formamide onto 5%, 4700 nucleotide pairs duplex, 5200 to 5700 total. It should be noted that the outer single-strand regions, *sl* and *sr*, also grow at the

Fig. 2 Partial homology maps of SPO2 /  $\phi$  105 heteroduplexes. (a) Heteroduplexes mounted from 50% formamide onto 20%; (b) heteroduplexes mounted from 40% formamide onto 10%; (c) from 30% formamide onto 5%; (d) histogram of the probability of being duplex for within the central region of homology for the sample of molecules mounted from 40% formamide onto 10%. The molecules were lined up by visual judgment in such a way as to give maximum matching of the duplex segments and of the single-strand loops between the different molecules. The length of the single-strand region to the left of the homology region is  $18,700 \pm 1000$  nucleotides long, or  $1.2 \times 10^7$  daltons duplex molecular weight. The single-strand to the right of the homology region has a length of  $15,600 \pm 800$  nucleotides, or  $1.0 \times 10^7$  daltons duplex molecular weight. In many of the substitution loops of the heteroduplexes, the two single strands appear to be of unequal length as indicated in the drawings. These length differences, in most cases in the 40% formamide spreadings, are less than 100 nucleotides long, and it is not certain whether the inequality is all due to experimental variability or whether some of it is real. The length measurements for  $\phi$  X 174 single-strand DNA gave the smallest standard deviation in the 40 to 10 % formamide spreadings. In the histogram in (d) one point was taken for every interval of 130 nucleotides starting from the left end of the first duplex segment for the 14 molecules in (b).

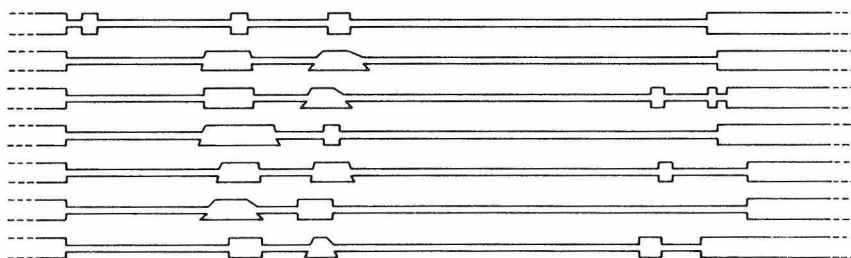
A. 50% formamide onto 20% Formamide



B. 40% formamide onto 10% Formamide



C. 30% formamide onto 5% Formamide



D.

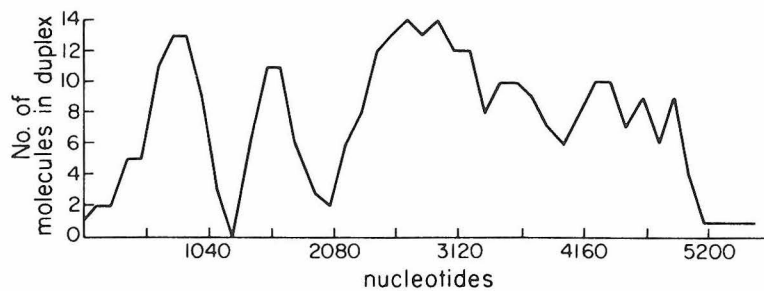


Fig. 2

expense of the duplex regions, *da* and *dd*, as the denaturing power of the solvent increases.

The heteroduplex patterns for molecules spread from 40% formamide onto 10% were studied most extensively. Fourteen molecules were traced and measured. The patterns of homology in the central segment were then lined up by visual judgment in such a way as to give what seemed to be maximum matching of the duplex segments and the single-strand loops between different molecules; these patterns are presented in Figure 2(b). A histogram of the probability of being duplex within the central region of homology, as derived from the plots in Figure 2(b), is presented in Figure 2(d).

#### 4. Further Discussion

The most important point is that there is a region of homology between the two phage DNA's. It is close to the center of the DNA molecules as obtained from phage particles and covers about 14% of the genome (5700 nucleotide pairs). At 30% formamide, this region is about 90% duplex. In 50% formamide, the fraction of the region which is duplex decreases to about 50%. This result strongly suggests partial homology between the two DNA's in this region.

The antigenic determinants responsible for serological inactivation of T2 and T4 bacteriophage are situated principally in the tail fiber and tail sheath regions (Franklin, 1961; Yanagida & Ahmad-Zadeh, 1970). It is probable that the antigenic determinants for serological inactivation of SPO2 and  $\phi$ 105 are similarly located in tail structures. Since the two phages show serological cross-reactivity and since their DNA molecules are homologous only in the central region, we conclude that the genes specifying some or all of the tail structures lie within this central region. This region is large enough to contain 8 to 10 typical size genes. It thus could include genes for a number of tail components and possibly for other phage functions.

A number of complementation groups for essential functions have been discovered by genetic studies for  $\phi$ 105 and SPO2 (Yasunaka, Tsukamoto, Okubo & Horiuchi, 1970; Armentrout & Rutberg, 1970). The genetic maps of both phages are linear, in agreement with our physical observation of no circular permutation. Many of the genetic markers are for unknown functions. The markers for DNA synthesis and for clear plaque formation in SPO2 (Yasunaka *et al.*, 1970) and in  $\phi$ 105 (Armentrout & Rutberg, 1970) are located away from the center of the genetic map. Markers for serum blocking power, and thus presumably for tail synthesis, have been observed to lie qualitatively in the center of the genetic map, with markers for DNA synthesis and for clear plaque formation one side and markers for head-related functions and unknown functions on the other side (W. R. Romig, personal communication; Farrell, 1970). Thus the available genetic data are qualitatively consistent with our conclusion that some or all of the genes for phage tail synthesis lie close to the center of the physical map. Furthermore, the combination of physical and genetic data suggests that the genetic sequences for clear plaque formation and for DNA synthesis of the two phages are non-homologous.

We now return to a consideration of the significance of partial homology between the DNA molecules of SPO2 and  $\phi$ 105. For the case of the heteroduplexes between the DNA molecules from phages 434,  $\lambda$  and 82 (and for most of the duplex segments between  $\lambda$  and 21), the patterns of duplex regions and single-strand loops are very reproducible and are not altered when the denaturing power of the spreading solution



and hypophase are varied by changing the formamide concentration at constant electrolyte concentration (Simon *et al.*, 1970). The duplex regions are apparently perfectly homologous. In the case of T7 and T3, however (Davis & Hyman, 1971), certain sequences are partially homologous—they appear as duplexes in spreading at low formamide concentration but open up to give single-strand regions at higher formamide concentration.

A quantitative discussion of the relation between the extent of sequence homology and the formamide concentration in which a duplex segment melts has been presented by Davis & Hyman (1971) and by Simon *et al.* (1971). Their arguments are based on measurements by McConaughy, Laird & McCarthy (1969) as to the effect of formamide concentration on the melting behavior of a duplex DNA. We thus estimate that for a duplex DNA with the average base composition of  $\phi 105$  and SPO2 DNA's (43.5% G + C), the conditions in the spreading solution of 0.10 M-cation concentration as used here, 25°C, and formamide concentrations of 50%, 40% and 30%, correspond to 22, 29 and 36 deg. C below  $T_m$ , respectively. From the data of Laird, McConaughy & McCarthy (1969) that 1.5% random base mismatches depress  $T_m$  by 1.0 deg. C, we then estimate that segments of the  $\phi 105$ /SPO2 heteroduplex that appear as duplex under these three conditions contain no more than 33, 43 and 54% partial sequence mismatch (assuming, of course, that their base composition is the average of 43% (G + C)). A duplex segment that opens up as the formamide concentration is raised from 30 to 40% formamide, for example, has a calculated degree of mismatch between 54 and 43%.

There are two non-homology regions which are always seen in 30% formamide (small loops at *sa* and *sb*) but each of these is only about 100 nucleotides long (see the top heteroduplex pattern in Fig. 4(c)); that is, less than the size of a gene. Much of the duplex region consists of partially homologous sequences, with very short regions, perhaps one or several bases, of base mismatch interspersed with equally short regions where the bases on the heteroduplex strands are complementary. The maps in Figure 2(a) suggest that there are no gene size regions (1000 nucleotides) which are always duplex; we do not know whether there are any regions of length greater than 100 nucleotides (for example) which are identical in the two phages.

In the case of the lambdoid phages (Simon *et al.*, 1971; Fiandt *et al.*, 1971), in general, gene size regions are either perfectly homologous or non-homologous. Apparently the high rate of genetic recombination between the phages maintains perfect homology in those sequences where recombination can occur. The same situation holds for T2 and T4 (Kim & Davidson, 1971). Partial homology throughout most of the genome is observed for T7 and T3 (Davis & Hyman, 1971). Genetic recombination between these two phages is rare; the partial homology is presumably due to evolutionary divergence from a common ancestor.

The DNA molecules of SPO2 and  $\phi 105$  are partially homologous in a central region comprising about 14% of the total molecular length and are non-homologous elsewhere. The partially homologous sequences presumably have resulted by evolutionary divergence from a common sequence. One possible explanation for the relation of the two phages is that they evolved from a common ancestor but the DNA sequences away from the center of the molecule have diverged so completely as to be non-homologous by the physical criteria used here. An alternative possibility for the relation of the two phages is suggested by the observation by Yamamoto (1969) that rare recombinants do form between the (apparently) completely unrelated bacterio-

phages P22 and Fels 2. Conceivably, the ancestors of  $\phi$ 105 and SPO2 were unrelated, but a rare recombination event between them gave rise to a hybrid phage containing the central genes of one and the outer genes of the other. Subsequent evolutionary divergence resulted in the partial homology within the central region. It may be that the study of homology relations for other phages related to SPO2 and  $\phi$ 105 will discriminate between these two explanations.

A point concerning the kinetics of renaturation deserves brief comment. If a heteroduplex is formed between equal concentrations of DNA molecules, AA' and BB', which are largely perfectly homologous, the percentage of heteroduplexes, is expected to be about 50 and this or just slightly less is about what is observed in heteroduplex studies in our laboratory. If the fraction of the genome that is homologous is  $f$  and if the rate of duplex formation is proportional to the number of nucleation sites available for initiating duplex formation, the ratio of heteroduplexes to homoduplexes should be  $f$ . For the 30 to 10% formamide experiments, this would predict about 10% of heteroduplexes in the present case. Partially homologous segments probably renature with each other more slowly than do perfectly homologous segments, although quantitative data are not available; thus, the fraction might be reduced to a predicted number of heteroduplexes of, say, 5%.

We actually observed 0.1 to 0.4% of heteroduplexes. The low rate of heteroduplex formation, we suggest, is due to the fact that the region of homology is close to the center of the molecule and is, according to the excluded volume effect (Wetmur & Davidson, 1968), less available for nucleating the formation of a duplex than is a homologous sequence closer to the topological outside of the molecule. It is our impression that the frequency of heteroduplex formation was higher for strands with broken ends than for intact strands, in agreement with the excluded volume hypothesis.

Another factor probably also operates to decrease the number of heteroduplexes on long annealing. If a strand A' mates with an available single-strand A region of an A · B' heteroduplex, it will then displace the B' strand by branch migration (Lee, Davis & Davidson, 1970), thus converting a heteroduplex to a homoduplex.

This research has been supported by the United States Public Health Service, grant GM 10991. One of us (L.B.B.) is grateful for the hospitality she has enjoyed in the laboratories of Dr E. S. Tessman and Dr E. K. Wagner.

#### REFERENCES

- Adams, M. H. (1959). *Bacteriophages*. New York: Interscience Publishers.  
 Armentrout, R. W. & Rutberg, L. (1970). *J. Virol.* **6**, 760.  
 Birdsall, D. C., Hathaway, G. M. & Rutberg, L. (1969). *J. Virol.* **4**, 264.  
 Boice, L. B. (1969). *J. Virol.* **4**, 47.  
 Boice, L. B., Eiserling, F. A. & Romig, W. R. (1969). *Biochem. Biophys. Res. Comm.* **34**, 398.  
 Bolton, E. P., Britton, R. J., Byers, T. J., Cowie, D. B., Hoyer, B., McCarthy, B. J., McQuillen, K. & Roberts, R. B. (1963). *Carnegie Institute Year Book*, **62**, 303.  
 Cowie, D. B. & Avery, R. J. (1970). *Carnegie Institute Year Book*, **69**, 528.  
 Davidson, N. & Szybalski, W. (1971). In *The Bacteriophage Lambda*, ed. by A. D. Hershey, ch. 3. New York: Cold Spring Harbor Laboratory.  
 Davis, R. W. & Davidson, N. (1968). *Proc. Nat. Acad. Sci., Wash.* **60**, 243.  
 Davis, R. W. & Hyman, R. W. (1971). *J. Mol. Biol.* **62**, 287.  
 Davis, R. W., Simon, M. N. & Davidson, N. (1971). In *Methods in Enzymology*, vol. 21 D, 413.



- Farrell, L. D. (1970). *Isolation and Characterization of Conditional Lethal Mutants of Bacteriophage SPO2*, Ph.D. Thesis, University of California, Los Angeles.
- Fiaandt, N., Hradečna, A., Lozeron, H. & Szybalski, W. (1971). In *The Bacteriophage Lambda*, ed. by A. D. Hershey, p. 329. New York : Cold Spring Harbor.
- Franklin, N. C. (1961). *Virology*, **14**, 417.
- Hausman, R. L., Magalhaes, E. P. A. & Araujo, C. (1961). *Anais de Microbiol. Brazil*, **9**, 511.
- Jacob, F. & Wollman, E. (1961). *Sexuality and the Genetics of Bacteria*. New York: Academic Press, New York.
- Laird, C. D., McConaughy, B. L. & McCarthy, B. J. (1969). *Nature*, **244**, 149.
- Lee, C. S. & Davidson, N. (1970). *Virology*, **40**, 102.
- Lee, C. S., Davis, R. W. & Davidson, N. (1970). *J. Mol. Biol.* **48**, 1.
- McConaughy, B. L., Laird, C. D. & McCarthy, B. J. (1969). *Biochemistry*, **8**, 3289.
- Naylor, H. B. & Burgi, E. (1956). *Virology*, **2**, 577.
- Romig, W. R. (1968). *Bact. Rev.* **32**, 349.
- Scalietti, J. V. & Naylor, H. B. (1959). *J. Bact.* **78**, 422.
- Schildkraut, C. L., Wierzechowski, K. L., Marmur, J., Green, E. M. & Doty, P. (1962). *Virology*, **18**, 43.
- Simon, M., Davis, R. & Davidson, N. (1971). In *The Bacteriophage Lambda*, ed. by A. D. Hershey, p. 313. New York: Cold Spring Harbor Laboratory.
- Summers, W. C. & Szybalski, W. (1968). *Virology*, **34**, 9.
- Westmoreland, B. C., Szybalski, W. & Ris, H. (1969). *Science*, **163**, 1343.
- Wetmur, J. G. & Davidson, N. (1968). *J. Mol. Biol.* **31**, 349.
- Wu, R. & Taylor, E. (1971). *J. Mol. Biol.* **57**, 491.
- Yanagida, M. & Ahmad-Zadeh, C. (1970). *J. Mol. Biol.* **51**, 411.
- Yamamoto, N. (1969). *Proc. Nat. Acad. Sci., Wash.* **62**, 63.
- Yasunaka, K., Tsukamoto, K., Okubo, S. & Horiuchi, T. (1970). *J. Virol.* **5**, 819.

## Part II

Electron Microscope Mapping of the Phage att Sites for  
Bacteriophages SPO2 and  $\phi$ 105, and of the Distribution  
of Ribosomal Genes on the Bacillus subtilis Chromosome

Louise T. Chow and Norman Davidson  
Divison of Chemistry and Chemical Engineering  
California Institute of Technology  
Pasadena, California, U.S.A. 91109

Electron Microscope Mapping of the Phage att Sites for Bacterio-  
phages SPO2 and  $\phi$ 105, and of the Distribution of Ribosomal Genes  
on the Bacillus subtilis Chromosome.\*

Louise T. Chow and Norman Davidson

Department of Chemistry  
California Institute of Technology  
Pasadena, California 91109

Summary

Circular duplex structures of the correct length are observed in the electron microscope in hybridization mixtures of lysogen DNA and mature phage DNA for the case of the temperate Bacillus subtilis bacteriophage SPO2. This result shows that the sequence order of the prophage is a circular permutation of that of the mature phage. By making heteroduplexes of prophage DNA with that of the SPO2 deletion mutants, R90 and S25, the att site of the phage has been mapped at  $61.2 \pm 0.6\%$  from one end of the mature phage DNA, which has a length of 38,600 base-pairs. In the same coordinate system, the R90 deletion extends from  $58.9 \pm 0.7$  to  $66.8 \pm 0.8\%$  on the SPO2 chromosome whereas the S25 deletion extends from  $63.2 \pm 0.6$  to  $66.9 \pm 0.7\%$ . In similar experiments with lysogen and mature phage DNA's of the temperate B. subtilis phage,  $\phi$ 105, no circular structures were seen. This result shows that the sequence order in the prophage and the phage are collinear, without circular permutation. Short

---

\*Contribution No. 4516

duplex segments, of length  $4830 \pm 250$  base pairs, with two single-strand arms at each end are seen at a low frequency after denaturation and renaturation of B. subtilis DNA. Several lines of evidence support the hypothesis that these duplex segments are formed by out-of-register renaturation of the 16S + 23S ribosomal RNA genes (rDNA) of B. subtilis. They are of the correct length. Their formation is inhibited if homologous but not if heterologous ribosomal RNA is added to the hybridization mixture. The frequency of occurrence of the duplex structures is consistent with the rDNA hypothesis. Heteroduplex molecules are seen with two or three rDNA duplex segments separated by single-strand substitution loops with specific lengths for each of the two single-strand arms of any one loop. On the basis of these structures, linkage groups containing 7 to 9 rDNA sets (each set containing one 16S and one 23S rDNA gene) separated by spacer DNA's are proposed. The evidence indicates that if 5S rDNA is present in the set it is located near one end to give a gene order 16S-23S-5S. All of the 16S rDNA genes are linked to 23S rDNA and vice versa with little or no spacer DNA between a 16S and 23S sequence. The spacer DNA between 23S and 5S must also be short. The prophage SPO2 bacterial att site maps at a distance 6200 bases away from a 16S + 23S rDNA set which is itself separated by a very short spacer (less than 600 bases) from a second rDNA set.

## 1. Introduction

$\phi$ 105 and SPO2 are temperate bacteriophages in their host Bacillus subtilis. The integration site of SPO2 is known to be near the erythromycin locus (ery-1) (Inselburg, Eremenko-Volpe, Greenwald, Meadow & Marmur, 1969) whereas  $\phi$ 105 integrates between the phe-1 and ilvA-1 markers (Rutberg, 1969).

In the Campbell model for lysogeny (Campbell, 1962), the infecting phage DNA cyclizes and is then inserted into the host chromosome by reciprocal recombination. The point on the phage DNA at which this reciprocal recombination occurs is called the att site. If the att site is separated from the site where the ends have joined (mn in Fig. 1) the DNA sequence order of the prophage will be a circular permutation of that of the mature phage. If, however, the att site is at the joint of the ends, the prophage sequence and the phage sequence will be collinear. These two alternatives are shown in Fig. 1. If the att site is very close to the joint of the ends, the prophage and phage sequences may appear to be collinear by all practical criteria.

By observing the structures of suitable heteroduplexes in the electron microscope, the position of the att site relative to the ends of the mature phage DNA can be physically mapped. This process is illustrated in Fig. 1. If a heteroduplex is formed between a strand of bacterial DNA containing the complete prophage and the complementary strand of the DNA of the mature phage, a circular duplex

Fig. 1 Scheme for mapping phage DNA attachment sites.

————, phage DNA; —————, prophage DNA.

m a b . . . . . n denote phage and prophage DNA sequences;

-----, bacterial DNA'.  $\alpha, \beta$  denote bacterial DNA sequences at  
at the site where the phage DNA integrates.

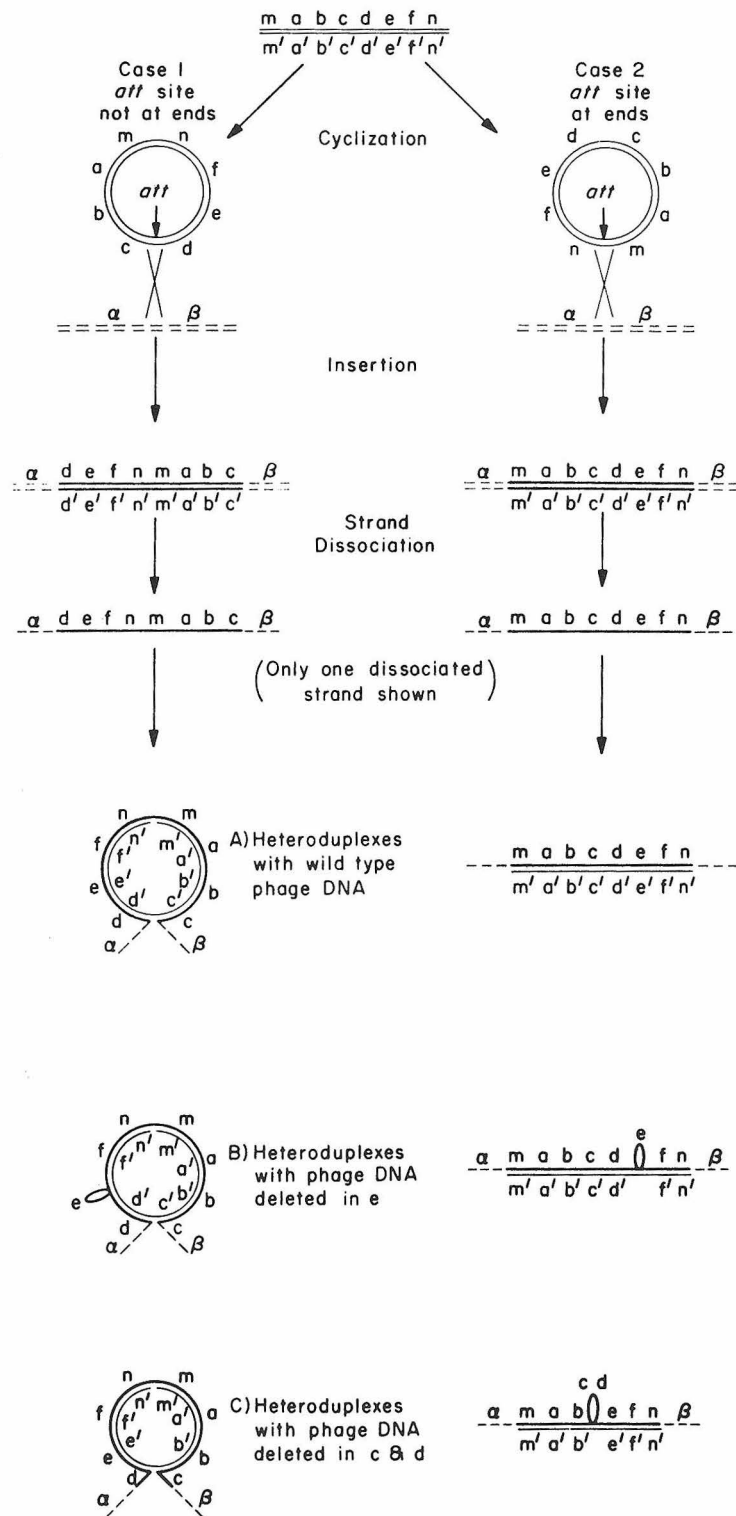


Fig. 1

structure will form when the att site is not at the ends, and a linear duplex structure will form if the att site is at the ends (Fig. 1A).

Figure 1 also shows the cases when the mature phage DNA used for the mapping is a deletion (or substitution) mutant of the prophage DNA (Fig. 1B and C). The nonhomology loop due to this deletion provides a reference point which makes it possible to map the att site for the case when a circular duplex has formed, and to confirm the mapping for a linear structure. In either case, observation of a structure with the deletion loop provides a strong confirmation of the identification of a duplex structure as being due to a prophage/phage heteroduplex. Such an experiment has been done with  $\lambda$  prophage DNA inserted into a bacterial episome (Sharp, Hsu & Davidson, 1972). The result of this work was to confirm the well-known map position of att <sup>$\lambda$</sup>  at 0.574 fractional  $\lambda$  lengths from the left end of the mature molecule. A similar experiment has been done by Hsu and Davidson (1972) on Mu-DNA inserted into an E. coli episome. In this case, the phage and prophage sequences are collinear. Episomes are convenient for such studies because they can be isolated away from the rest of the bacterial DNA, and contain a defined sequence of bacterial DNA. In the cases of B. subtilis lysogens, no episomes are available. We have therefore done the more difficult experiment of hybridizing the entire bacterial DNA with excess phage DNA and looking for the small fraction of duplexes of one or the other types depicted in Fig. 1.



It should be noted that genetic studies indicate that for the limited number of markers available the marker order in prophage  $\phi 105$  and in phage  $\phi 105$  are collinear (Armentrout & Rutberg, 1970). For SPO2, the marker order of the prophage is circularly permuted relative to that of the mature phage (S. Truesdell, E. Scibienski, & W. R. Romig, private communication).

In the course of these studies we have observed in the renatured bacterial DNA, duplex structures which we interpret as being due to renaturation between out-of-register ribosomal RNA sequences of B. subtilis. These studies provide information about the distribution of the ribosomal DNA sequences along the B. subtilis 168 chromosome.

## 2. Materials and Methods

Bacterial strains, phages and phage DNAs. The lysogens Bacillus subtilis 168 (SPO2) and Bacillus subtilis 168 ( $\phi$ 105) were gifts from L. B. Boice, and were isogenic. S25 phage, a deletion mutant of SPO2, and a mixture of the DNA's of SPO2 wild-type phage and the deletion mutant R90 were gifts from R. Skillern. Phage  $\phi$ 105 was a gift from L. B. Boice.

Media. TY broth contained 10 gm of tryptone (Difco), 5 gm of yeast extract (Difco), and 10 gm of NaCl per liter of water. The pH was adjusted to 7.4 with NaOH before autoclaving. 0.6 ml of 0.5 M  $\text{CaCl}_2$  and 1 ml of 0.01 M  $\text{MnCl}_2$  were added per liter of medium after sterilization. Plating agar contained 1.5% agar supplemented with TY broth at pH 6.8.

Preparation of lysogen DNA's. B. subtilis 168 (SPO2) cultures were grown with aeration at 37°C in TY broth by inoculating either directly from a slant or from a single colony from an agar plate. B. subtilis 168 ( $\phi$ 105) was grown in TY broth from a slant. Growth of a 250 ml culture was stopped when  $A_{600}$  reached between 0.4 and 0.5. Cells were spun down at 4°C in a Sorvall GSA rotor, washed twice in TES buffer (0.05 M Tris, 0.005 M EDTA, 0.05 M NaCl, pH 8), resuspended in 4 ml of TES containing 1 mg/ml lysozyme and 0.4 g of sucrose, and then incubated at 37°C for 10 min. Lysis was accomplished by adding 2 ml of 2% sarcosyl. The viscous lysate was sheared 5 times in a 10 ml plastic disposable syringe at

a speed of 5 to 10 sec/volume. The solution density was adjusted to 1.7 g/ml by adding saturated CsCl solution in 0.1 M Tris, 0.01 M EDTA, pH 8, and solid CsCl. DNA banding was carried out for 43 hr in an SW 50.1 rotor at 37 krpm at 20°C, with a Beckman L2-65B ultracentrifuge. The gradient was dripped from the bottom of the centrifuge tube. Fractions containing DNA were obviously more viscous. They were pooled and stored as such. The length of the DNA thus prepared had a single strand length ranging from about 30  $\mu$  to 60  $\mu$ , with a few longer strands.

Preparation of B. subtilis ribosomal RNA. A three liter culture of B. subtilis 168 ( $\phi$ 105) was grown in TY broth. When  $A_{600}$  reached 0.5, the cells were spun down at 4°C and resuspended in 18 ml of 25% sucrose in 0.04 M Tris, pH 8. 4.8 mg of Bentonite (Sigma) and 1.8 ml each of lysozyme (6.4 mg/ml in 0.25 M Tris, pH 8) and EDTA (20 mg/ml) were added. The suspension was allowed to warm to room temperature. The cells were lysed by adding 1.35 ml each of 0.1 M  $MgSO_4$  and 5.2% Brij58 followed by 4 cycles of freezing (in a -70°C freezer) and thawing. Cell debris and DNA were removed by spinning in an SW 50.1 rotor at 43 krpm for 10 min. Ribosomes were pelleted by centrifugation in an SW 50.1 rotor for 2.5 hr at 43 krpm at 5°C. The pellet was resuspended in 2 ml of cold TMM (0.002 M mercaptoethanol, 0.005 M Tris, 0.01 M  $MgSO_4$ , pH 7.2). 20  $\mu$ l of 1 mg/ml DNase (Worthington DPFF) were added and the suspension was allowed to stand on ice for 15 min.

Ribosomes were again pelleted from TMM at 43 krpm at 4°C for 2.5 hr in an SW 50.1 rotor, resuspended in 1.5 ml of cold TMM, and stored in a -70°C freezer.

40  $\mu$ l of 17% SDS was added to 200  $\mu$ l of ribosome solution ( $A_{260} = 85$ ). This solution was layered onto two 5 to 30% sucrose gradients in 0.5% SDS, 0.1 M NaCl, 0.01 M Tris, 0.001 M EDTA, pH 7.2. After centrifugation at 44 krpm at 20°C for 4 hr in an SW 50.1 rotor, 25 drop fractions were collected from the bottom of the centrifuge tubes.  $A_{260}$  was read to locate the RNA's. The peak fractions had  $A_{260}/A_{280} = 2.0$ . 95% of the material sedimented in the 23S and 16S bands. To the 23S and 16S fractions were separately added 2.5 volumes of cold ethanol; the solutions were stored at -70°C overnight. RNA precipitates was spin down at 10 k for 15 min in a Sorvall SS-34 rotor at 4°C. The supernatant was discarded and the pellet was resuspended in 2 ml of 0.3 M sodium acetate, pH 5. Ethanol precipitation was repeated. The RNA's were recovered in 400  $\mu$ l of 0.1 M sodium acetate, pH 5, and stored in a -70°C freezer.

HeLa 28S RNA, extracted from total HeLa cells, was a gift from B. Attardi.

### Electron Microscopy

(a) DNA-DNA hybridization. DNA heteroduplexes between phages were prepared as described before (Chow, Boice & Davidson, 1972). To prepare heteroduplexes between phage and lysogen DNA's, denaturation with NaOH and reneutralization was carried out as before. Annealing was performed by dialysing against 0.1 M Tris,

0.01 M EDTA, 0.25 M NaCl, pH 8.5 in 70% formamide, for 2 hr at room temperature. Renaturation was stopped by dialysing against 0.01 M Tris, 0.001 M EDTA, pH 8 at 4°C. In self-renatured preparations, only one kind of DNA was used. The DNA concentrations for the various heteroduplex preparations were as follows: (1) phage/phage DNA heteroduplexes, 2.5  $\mu\text{g/ml}$  of each kind; (2) phage/lysogen heteroduplexes, 10  $\mu\text{g/ml}$  of lysogen DNA, 0.5  $\mu\text{g/ml}$  of phage DNA.

(b) Hybridization with RNA present. Lysogen DNA or a mixture of lysogen DNA and phage DNA was denatured and reneutralized as described before. RNA was then added at a concentration of 10  $\mu\text{g/ml}$ . When B. subtilis ribosomal RNA's were used, 23S and 16S RNA were added at a weight ratio of 2 to 1. Renaturation was carried out as described above. The solution was then treated with pancreatic RNase (20  $\mu\text{g/ml}$ ) at room temperature for 30 min (pancreatic RNase stock solution was made in 0.1 M sodium acetate buffer, pH 5, and heated to 90°C for 10 min to inactivate any contaminating DNase). The bulk of the degraded ribonucleotides was removed by first dialysing against 0.1 M NaCl, 0.01 M Tris, 0.001 M EDTA, pH 8, overnight and then against 0.01 M Tris, 0.001 M EDTA, pH 8, at 4°C.

Snap back of denatured lysogen DNA. Lysogen DNA was denatured with NaOH and reneutralized as described before. The denatured DNA was either dialysed against 0.01 M Tris, 0.001 M EDTA, pH 8 at 4°C immediately or was dialysed against 0.1 M Tris, 0.01 M EDTA and 60% formamide, pH 8.5, at 4°C for 2 hr and then

against 0.01 M Tris, 0.001 M EDTA, pH 8, at 4°C.

Preparation of grids. The spreading solution for electron microscopy contained 0.076 M Tris-OH, 0.024 M Tris-HCl, 0.01 M Na<sub>3</sub>EDTA, pH 8.5, 50  $\mu$ g/ml cytochrome c, 1-2  $\mu$ g/ml DNA when lysogen DNA was used and 0.5  $\mu$ g/ml of DNA when only phage DNA was used, and 45% (by volume) formamide; 50  $\mu$ l were used per spreading. The hypophase contained 1/10 of the electrolyte concentrations and 17% formamide.

200 Mesh copper grids supported by parlodion films were used. Grids were stained with uranyl acetate and then rotary-shadowed with Pt-Pd alloy.  $\phi$ X174 DNA and  $\phi$ X174 RF II DNA were mounted on the same grids as single-strand and duplex DNA length standards, with molecular weights taken as  $1.72 \times 10^6$  and  $3.45 \times 10^6$  daltons or 5200 bases and base pairs, respectively.

### 3. Results

#### (a) Structures of SPO2 and $\phi$ 105 Prophages

Mapping of SPO2 deletion mutants. As indicated in Fig. 1, it is helpful to do the prophage/phage heteroduplex experiments with a phage DNA carrying a deletion that serves as a reference point along the phage genome. Mr. R. Skillern in Professor W. R. Romig's laboratory at the University of California at Los Angeles has isolated a number of SPO2 deletion mutants that have a clear plaque morphology on plating bacteria and that cannot complement each other. They are believed to be deleted in phage functions essential for lysogeny. Two of these phages, S25 and R90, were generously given to us by R. Skillern.

The heteroduplexes between wild-type SPO2 and between the two deletion mutants were prepared as described in Materials and Methods. Illustrative micrographs are presented in Plate I; the mapping data are presented in Fig. 2.

R90 was known from the work of Skillern to be less sensitive to pyrophosphate inactivation than S25 and therefore most probably to contain a larger deletion. The mapping experiments show that the DNA is deleted in  $7.9 \pm 0.4\%$  of the SPO2 genome, whereas S25 DNA is deleted in  $3.7 \pm 0.4\%$  of the SPO2 genome. The right end points (with the arbitrary assignment of left and right given in Fig. 2) of the deletions in the two DNAs are the same within experimental error.

Plate I    Heteroduplexes between DNA's from wild type and the deletion mutants, R90 and S25, SPO2 phages.

A   R90 / S25 heteroduplex.

B   wild type / R90 .

C   wild type / S25 .

Heteroduplexes were prepared as described in the text, and were mounted from 45 % formamide onto 17 % formamide. The small circular molecules are  $\phi$  X 174 and  $\phi$  X 174 RFII.



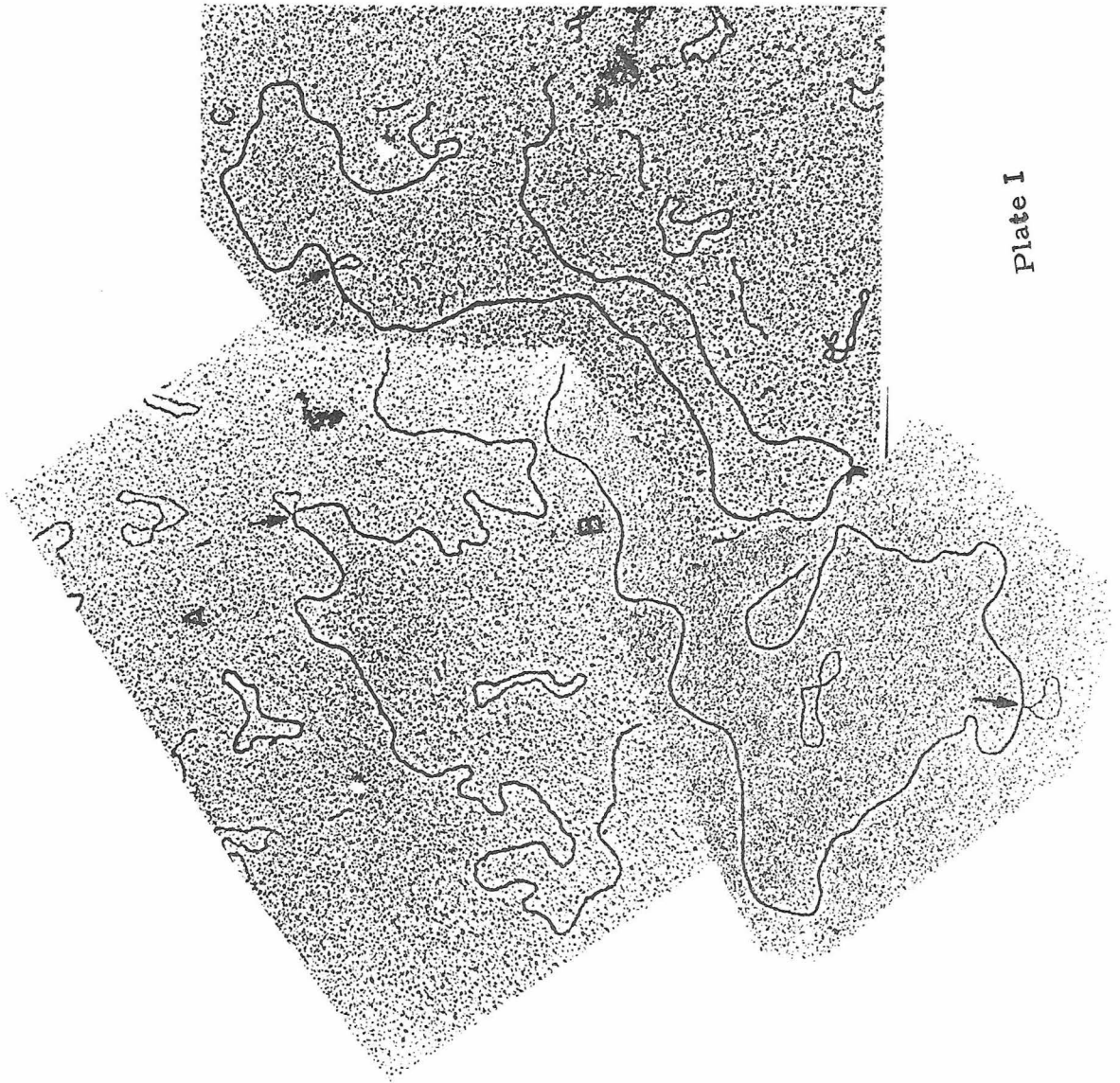


Plate I

Fig. 2    Physical map of the DNA of the SPO2 deletion mutants, R90 and S25.

Heteroduplexes with wild type SPO2 , R90 and S25 DNA's were prepared as described in Materials and Methods. The figure also shows the position of the att site as determined from prophage / phage heteroduplexes. The left and right directions on the molecule are oriented arbitrarily. The unit kb is kilobases; that is , 1000 bases or base pairs.

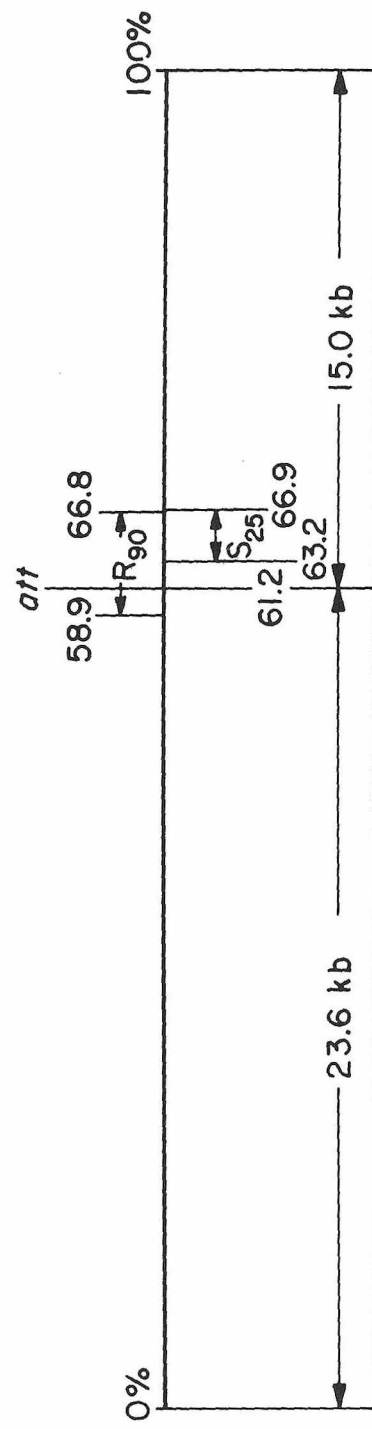


Fig. 2

In the heteroduplex R90/S25, a deletion loop of length  $4.0 \pm 0.3\%$  of the SPO2 genome was seen rather than a substitution loop. If the S25 deletion is entirely within the R90 deletion, a deletion loop will be seen for such a heteroduplex. If the right end of the S25 deletion extends beyond the right end of the R90 deletion, a substitution loop will be seen. The estimated error in mapping the positions of the deletion loops in the heteroduplexes with wild-type phage is  $\pm 150$  nucleotides. The accuracy in distinguishing a deletion loop from a substitution loop is probably 50 nucleotides. We conclude therefore that the right end of the S25 deletion is less than 150 nucleotides to the left of and less than 50 nucleotides to the right of the right end of the R90 deletion. The two end points may be identical at a hot spot for illegitimate recombination.

Mapping of the att site of phage SPO2. Heteroduplexes were prepared by renaturing samples of rather high molecular weight bacterial DNA prepared from an SPO2 lysogen (B. subtilis 168 (SPO2)) with a mixture of wild-type and R90 phage DNAs. Much of the DNA seen was either single- or double-strand renatured bacterial DNA and phage DNA. A low frequency of circular duplex structures with single-strand branches, as indicated in Fig. 3, were seen. The circular duplex structures corresponding to heteroduplexes between prophage and phage DNA occurred at a frequency of about 0.7 per grid square examined. No circular structures were seen in self-renatured B. subtilis 168 (SPO2) DNA without added SPO2 DNA or in self-renatured SPO2 DNA by itself. (It should be recalled (Chow et al.,

Fig. 3 Schematic representation of typical heteroduplexes observed between prophage ( B. subtilis 168 (SPO2) ) and SPO2 wild type, R90, or S25 DNA's.

- (a) Intact prophage / wild type heteroduplex.
- (b) Intact prophage / R90 or wild type heteroduplex.
- (c) Broken prophage / intact wild type or R90.
- (d) Intact prophage / broken wild type or R90.
- (e) Intact prophage / two broken wild type or R90.
- (f) Intact prophage / S25.
- (g) and (h) Broken prophage / intact S25.
- (i) Two pieces of broken prophage / intact S25.
- (j) Intact porphage / two pieces of broken S25.

As indicated by the dotted line, some structure of type j were seen without a gap. These are due to hybrids with two broken phage DNA strands, each of which has one intact end. In structures a , b , c , e , f , g , h , and i , the position of the ends of the mature phage DNA in the circular heteroduplex is indicated by a gap.

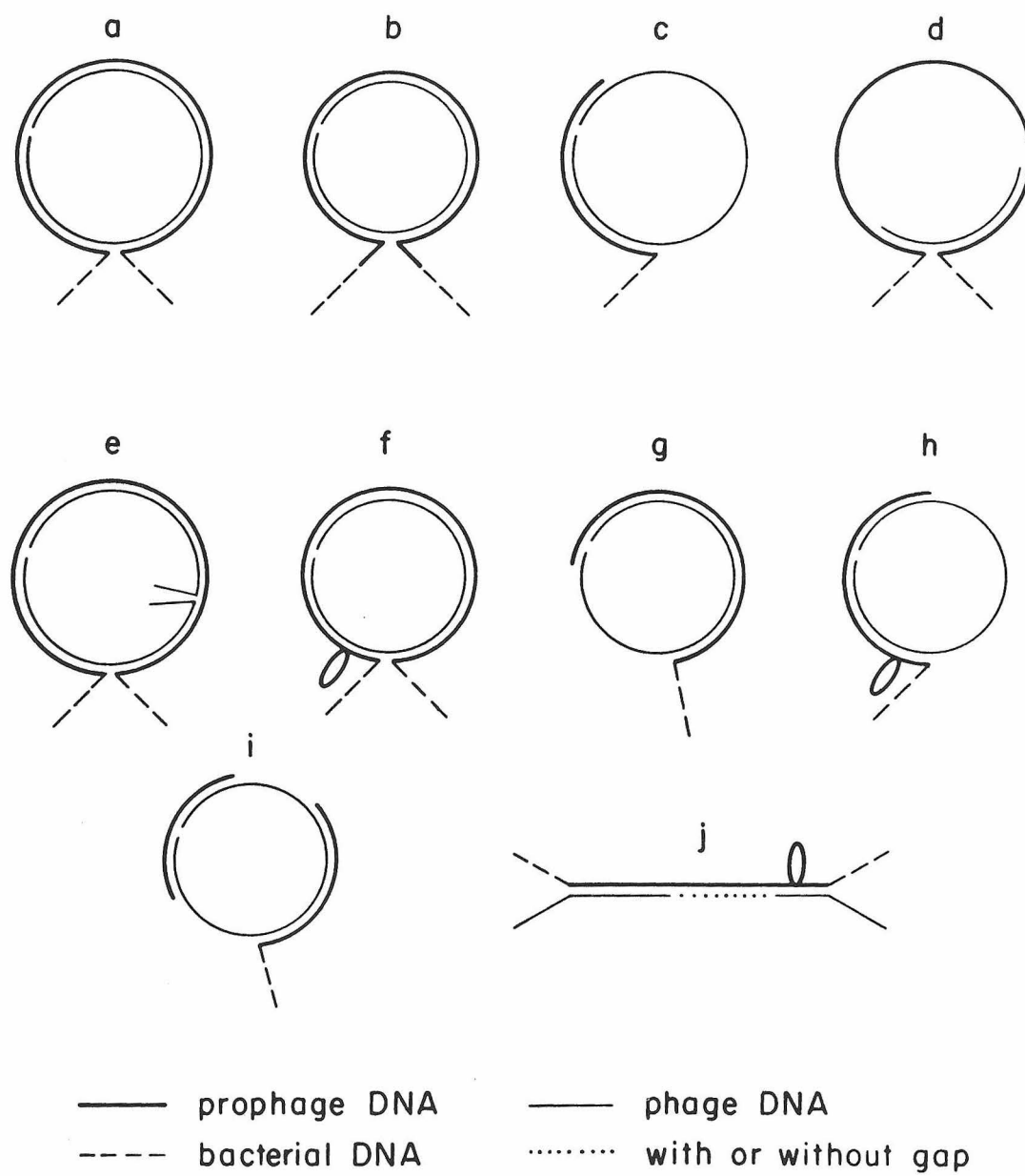


Fig. 3

1972) that SPO2 and  $\phi$ 105 DNA's each has a unique non-permuted sequence and cohesive ends. Each can cyclize, but the hydrogen bonded circular structures are not stable under the formamide spreading conditions used here. )

The duplex circles were of two sizes corresponding to wild-type and R90 DNAs. The results are presented in Table 1. Not all of the circular duplexes were of the ideal type indicated in Fig. 3A. Some of them contained single-strand gaps or extra branches resulting from pairing of one or two broken phage DNAs with prophage DNA and vice versa. These structures are also depicted in Fig. 3. Typical micrographs are presented in Plate II.

The observation of a circular duplex shows that the prophage sequences are a circular permutation of the phage DNA sequences. The R90 deletion loop is not seen in the duplex structures. We conclude that the att site is located within the sequence deleted in R90. This is therefore an example of the kind of duplex depicted in Fig. 1C.

Heteroduplexes between S25 DNA and B. subtilis 168 (SPO2) DNAs were then prepared and examined. In this case, duplex circular molecules with the deletion loop were indeed found (Plate II). The position of the deletion loop is about 800 base pairs from the att site. Since the att site is within the R90 deletion, we conclude that it is located to the left of the left terminus of the S25 deletion and maps as indicated in Fig. 2. The observation of heteroduplexes with this

TABLE 1

Lengths of circular structures\* in hybridization of B. subtilis 168  
(SPO2) DNA with SPO2 wild type R90 DNA\*\*

No. of Circular Structures Observed	Length (kb)	Designation	Expected <sup>†</sup> Length (kb)
9	$38.5 \pm 0.5$	Heteroduplex of prophage and wild type DNA	38.6
6	$35.9 \pm 0.5$	Heteroduplexes of prophage and R90 DNA	35.5

\* Complete and partial duplex circular structures as in  
Fig. 3 a, b, c, d, and e.

\*\* Hybridization was carried out as described in Materials and  
Methods.

<sup>†</sup> Expected values from length measurements on phage  
heteroduplexes.



Plate II SPO2 prophage ( *B. subtilis* 168 (SPO2) ) / SPO2 phage heteroduplexes.

- A Linear heteroduplex between wild type prophage and S25 DNA's; arrows point to phage attachment sites, the S25 deletion loop, and a single-stranded gap corresponding to one or two broken ends of the two mature phage DNA molecules ( see Fig. 3 j ).
- B Circular heteroduplex between prophage and S25 DNA's; arrows point to the phage attachment site and the S25 deletion loop ( see Fig. 3 f ).
- C Circular heteroduplex between prophage and R90 DNA's; arrow points to the phage attachment site ( see Fig. 3 b ).

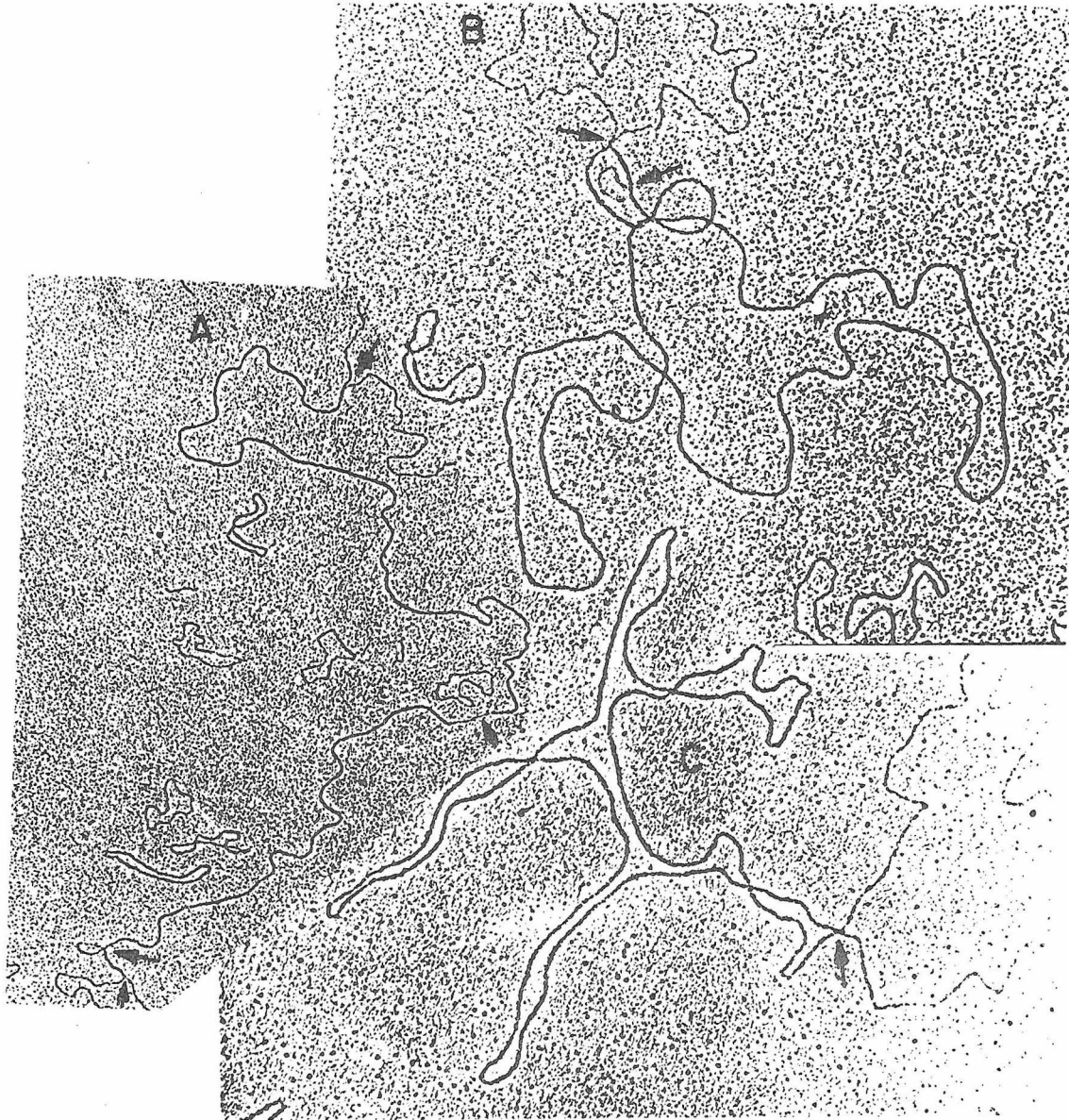


Plate II

unique deletion loop, of course, gives one greater confidence that the overall interpretations are sound and the results are not due to some peculiar artifact.

As a technical matter, it should be noted that some linear heteroduplexes resulting from pairing of prophage and two broken pieces of phage DNAs were also seen. A typical structure is depicted in Fig. 3. These can be identified with more certainty in the S25 case because the S25 deletion loop is seen in the duplex region.

$\phi$ 105 prophage is not circularly permuted. As noted above, the frequency of occurrence of circular heteroduplexes between SPO2 DNA and its prophage is about 0.7 per grid square with the DNA concentrations and conditions of mounting given. When heteroduplexes between  $\phi$ 105 DNA and B. subtilis 168 ( $\phi$ 105) DNAs were prepared under identical conditions and examined, no circular duplexes were seen in an examination of approximately 40 grid squares. We conclude that  $\phi$ 105 probably inserts into its host DNA without circular permutation. In a similar case of Mu prophage DNA/Mu DNA heteroduplexes, Hsu and Davidson (1972) were able to recognize the linear duplex region because the entire structure was contained within a circular episome and because the Mu DNA itself contains certain nonhomology loop features that aid identification. In the present case, no deletion mutants of  $\phi$ 105 were available. Linear single strands and linear duplex DNA due to renaturation of the host B. subtilis DNA itself were present in large quantities. It was therefore

impossible to positively recognize a linear duplex region due to a prophage  $\phi 105$  heteroduplex. Thus, our conclusion about the collinearity of prophage and phage  $\phi 105$  DNA is based entirely on the absence of any circular structures.

(b) Mapping Repetitious (ribosomal?) Sequences on *B. subtilis* DNA

Self-renaturation of *B. subtilis* DNA. There are approximately ten 16S and ten 23S rRNA genes in the *B. subtilis* chromosome of molecular weight  $3.9 \times 10^9$  (Smith, Dubnau, Morell & Marmur, 1968; Eberle & Lark, 1967). There is evidence that a 16S and 23S sequence are very closely linked (Colli & Oishi, 1969). In some cases at least, there is a 5S RNA sequence attached on the 23S side (Colli, Smith & Oishi, 1971). Dolittle and Pace (1971) have offered evidence that 16S, 23S, and 5S ribosomal RNAs of *E. coli* are transcribed in vivo as a single transcriptional unit in the order 16S-23S-5S. If the same is true of *B. subtilis*, the total molecular length of a 16S plus 23S ribosomal DNA set would be  $1.65 \times 10^6$  daltons or 4.8 kb; the length would be slightly greater if there were a small spacer between the 16S and 23S sequences and/or if a 5S sequence were also present.

If there were no repeated sequences in a sample of DNA, then on denaturation the only duplex structures seen would be linear duplexes with one or no single strands at each end. If a particular sequence is repeated in the chromosome, renaturation of the out-of-register sequences would give the kind of structure depicted in

Fig. 4o with, in the ideal case, two single-strand arms emerging from each end of the duplex.

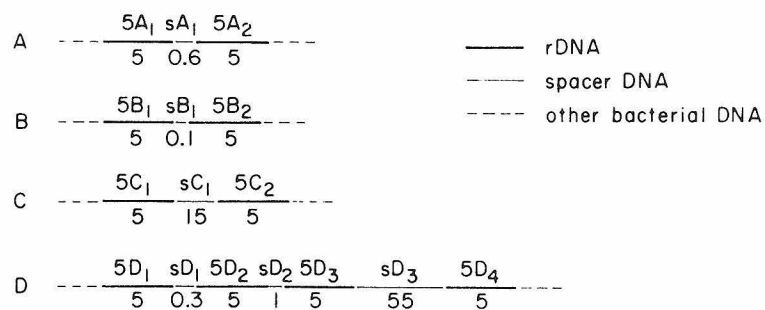
While scanning grids in search of the prophage/phage hetero-duplexes and later by scanning grids prepared by self-renaturing B. subtilis DNA with no phage DNA present, we have seen the various structures depicted in Fig. 4. Illustrative electron micrographs are presented in Plates III and IV. The basic unit of the structure is a short duplex region of length about 5000 base pairs. In typical cases, there are two single-strand arms at each end. In many cases, there are two duplex regions. They are connected by a loop consisting of two single-strand arms, and with two additional single-strand arms at each outside end of the duplexes. Several structures (Fig. 4j, m, n) with three duplex regions have been seen. We shall refer to structures with one, two, and three duplex regions as singlets, doublets, and triplets.

The detailed features of the several different kinds of structures seen and their interpretation in terms of the distribution of ribosomal DNA sequences will be discussed later. We first describe experiments to test the hypothesis that the duplexes are due to renaturation of ribosomal DNA sequences. If the short duplex structures observed are indeed due to renaturation of ribosomal DNA sequences, their frequency of occurrence should be greatly reduced if the renaturation takes place in the presence of excess Bacillus subtilis rRNA but should not be affected by a heterologous RNA.

Fig. 4 Simple rDNA linkage groups, and rDNA heteroduplex structures observed.

The numbers below the representations of the linkage groups are the lengths of the respective segments in kb. Particular sequences are identified by symbols above the linkage groups; for example, 5A<sub>1</sub>, and 5A<sub>2</sub> are the first and second rDNA sequences of linkage group A; sD<sub>2</sub> signifies the second spacer in linkage group D. Length measurements and standard deviations of the spacers are reported in Table 4. Spacers sB<sub>1</sub> and sD<sub>1</sub> were measured as  $0.14 \pm 0.05$  and  $0.28 \pm 0.05$  kb, respectively. For identification purposes, they are labeled 0.1 and 0.3kb. Further discussion about sB<sub>1</sub> and sD<sub>1</sub> is presented in the text. The observed heteroduplex structures are shown in the lower part of the figure. They were observed in self-annealed B. subtilis 168 (  $\phi$  105 )( not clones ), self-annealed B. subtilis 168 ( SPO2 ) (not cloned, and cloned ), and hybridization preparations between B. subtilis 168 ( SPO2 ) and SPO2 wild type plus R90 or S25 DNA's. In the heteroduplex structures, the lengths of the spacers are given, but the duplex lengths are not. Below each structure, its interpretation as a heteroduplex of two of the linkage groups is shown; the number of molecules observed is also given. \* We believe that in this heteroduplex, due to the unusual spiral structure, the large spacer was subjected to an abnormal stretching force, and was therefore extended from 55 to 66 kb. \*\* This number includes 7 singlets seen adjacent to SPO2 prophage / phage heteroduplexes.

## Linkage groups



## Structures

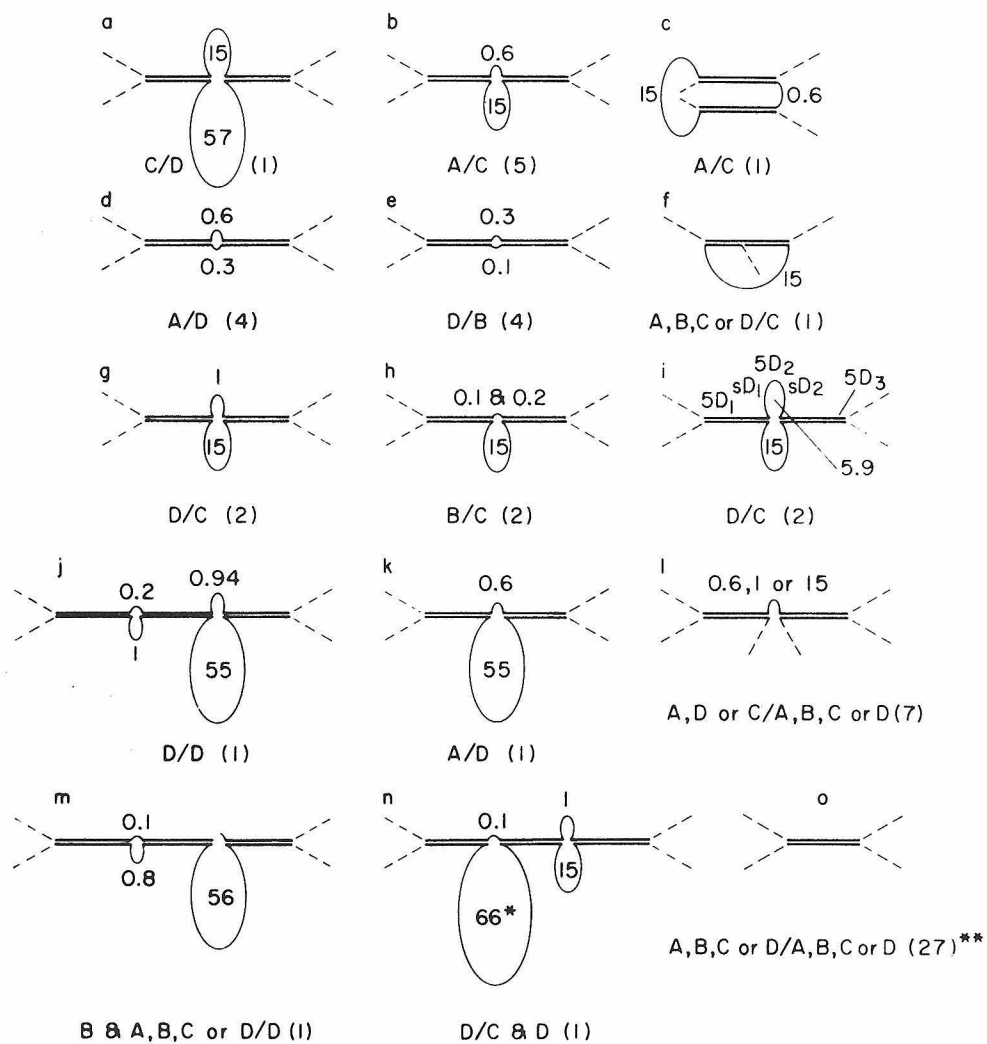


Fig. 4

Plate III Out-of register ribosomal DNA heteroduplexes from  
B. subtilis 168 ( $\phi$  105) and from B. subtilis 168 (SPO2).

- A Doublet of the type depicted in Fig. 4d or 4e.
- B Fig. 4i type doublet .
- C Fig. 4g type doublet; an interpretative tracing on a reduced scale is also shown.
- D Fig. 4 c type doublet.

The DNA's were denatured and renatured in the presence or absence of SPO2 phage DNA, as described in the text. Arrows point to single-strand spacer DNA segments and to the junctions of duplex rDNA and single-strand DNA arms. Some particular segments of the linkage groups at the top of Fig. 4 are identified on the micrographs. ———, rDNA ; ———, sC<sub>1</sub> spacer DNA; - - - - - , sD<sub>2</sub> spacer DNA.



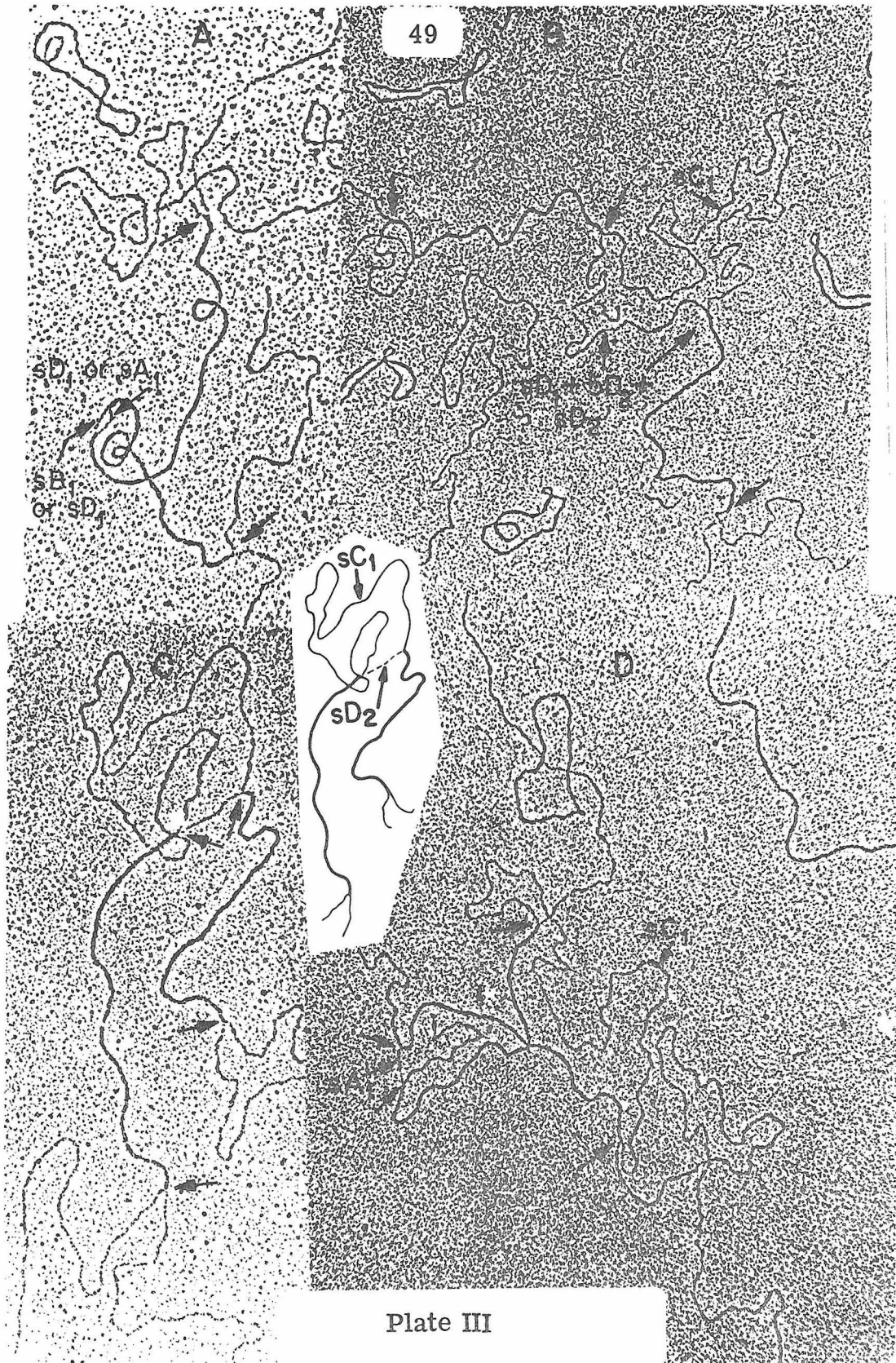


Plate III

Plate IV Out-of-register ribosomal DNA heteroduplexes.

- A Doublet of the type depicted in Fig. 4a, with interpretative tracing.
- B Triplet as in Fig. 4j; - - -> indicates an uncertain junction between rDNA duplex and the single-strand arms.

For other details and notation see Plate III and Fig. 4.

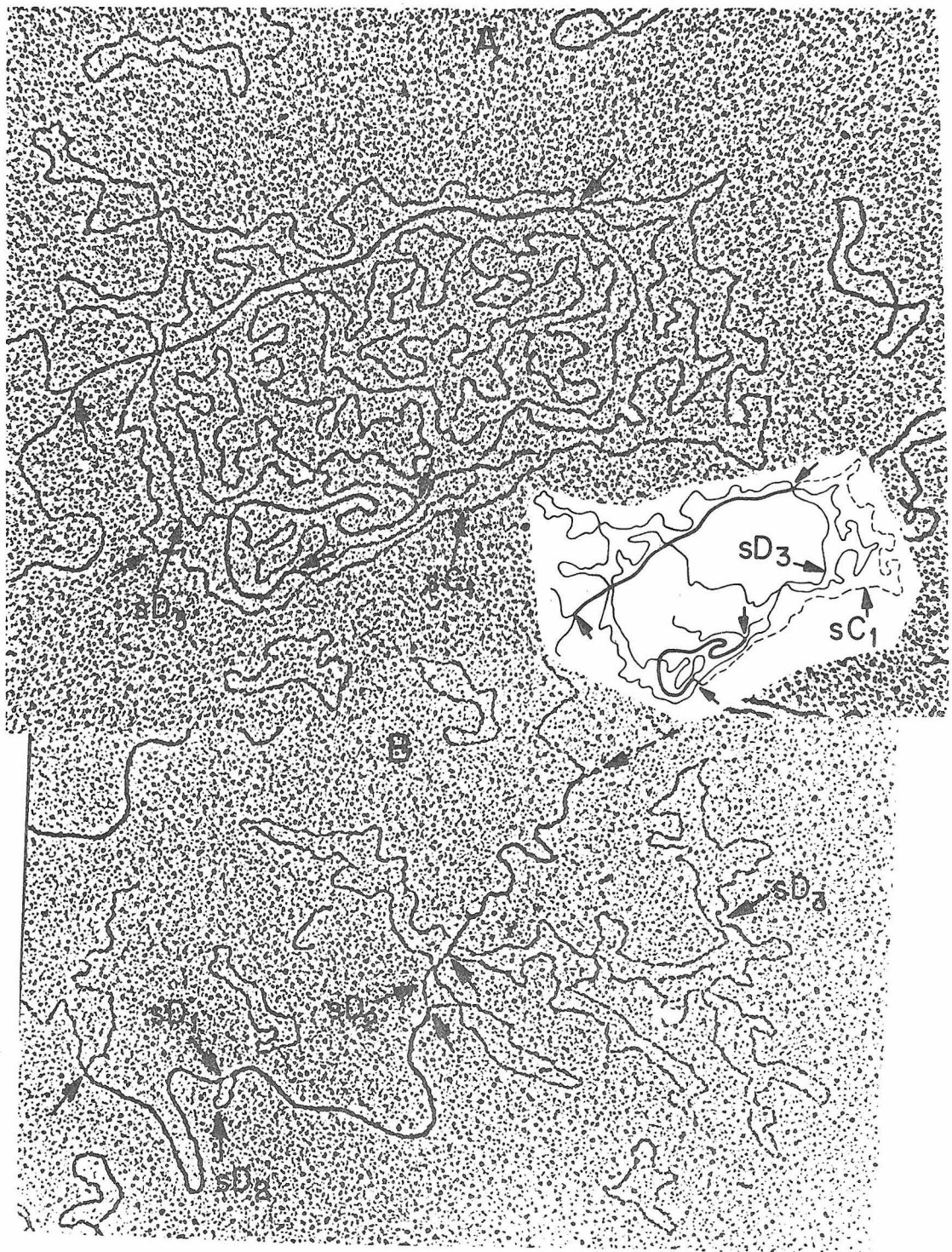


Plate IV

For the latter RNA we have used HeLa 28S RNA.

16S and 23S B. subtilis rRNA were included in a renaturation mixture of B. subtilis 168 (SPO2) DNA as described in Materials and Methods; in control experiments HeLa 28S RNA was used. The results are presented in Table 2. In the absence of any added RNA, the duplex structures attributed to rDNA occur at a frequency of about 0.5 to 0.75 per grid square. The data in Table 2 show that although HeLa 28S rRNA does not significantly affect this frequency, the formation of the duplex structures we identify as rDNA sequences is greatly reduced if Bacillus subtilis rRNA is present in the incubation mixture. The data also show that B. subtilis rRNA does not inhibit the formation of prophage/phage heteroduplexes. We conclude that the duplex structures seen are indeed due to out-of-register renaturation of rDNA sequences.

We may ask whether the observed frequency of the out-of-register duplexes is consistent with the rDNA hypothesis.

The total amount of single-strand and duplex DNA on randomly chosen areas of several grid squares was measured. These data are reported in Table 3. We find that under the renaturation and spreading conditions used, there are about 2.75 ( $\pm 0.25$ ) B. subtilis chromosomes per grid square, with 25% or 0.70 chromosomes present as long in-register duplexes. If there are ten 16S plus 23S rDNA sets per chromosome (Smith et al., 1968), there should be 7 rDNA duplex sequences per grid square present as parts of the long renatured in-register duplexes. The average length of a single strand in the

TABLE 2

Effect of RNA on frequency of out-of-register duplex structures\*

DNA	RNA	No. of rDNA Duplexes** Observed	No. of Grid Squares Scanned	No. of Circular SPO2 Prophage/phage Heteroduplexes
<u>B. subtilis</u> 168 (SPO2)	HeLa 28S	23	43	0
<u>B. subtilis</u> 168 (SPO2)	B. subtilis ribosomal 16S + 23S	1	15	0
<u>B. subtilis</u> 168 (SPO2), SPO2 wild type, R90	B. subtilis ribosomal 16S + 23S	1	14	8
<u>B. subtilis</u> 168 (SPO2), S25	B. subtilis ribosomal 16S + 23S	0	20	13

\* Hybridization of B. subtilis 168 (SPO2) DNA with phage DNA and RNA was carried out as described in Material and Methods.

\*\* Duplex structures of length  $4.83 \pm 0.25$  kb with two single-strand arms at each end.



TABLE 3

Amount of DNA on a 200-mesh electron microscope grid square\*

Area**	Amount of*** Duplex DNA $\times 10^6$ daltons	Amount of*** Denatured DNA $\times 10^6$ daltons	Total Amount of DNA per Area $\times 10^6$ daltons	Total Amount† DNA per Grid Square $\times 10^9$ daltons	% of DNA Renatured
1	44.0	119.5	163.5	12.1	27
2	41.3	93.6	134.9	10.0	31
3	39.7	102.3	142.0	10.5	28
4	25.0	114.3	139.2	10.3	18
Ave.				$10.7 \pm 0.9$	$26 \pm 6$
				or $2.75 \pm 0.25$ B. subtilis chromosomes‡	

54

\* Under the renaturation and spreading conditions described in Materials and Methods.

\*\* Randomly chosen areas were photographed.

\*\*\* All duplex and denatured DNA strands present on the photograph were measured.

† One grid square of a 200-mesh grid was measured to have an area of about  $6.4 \times 10^{-5}$  cm<sup>2</sup>. Each photograph (magnification 3000) covered an area of  $8.7 \times 10^{-7}$  cm<sup>2</sup> or 1/74 of a grid square.‡ The molecular weight of one B. subtilis chromosome is taken as  $3.9 \times 10^9$  daltons (Eberle and Lark, 1967).

preparation is about  $5 \times 10^7$  daltons; the length of an rDNA set is  $1.5 \times 10^6$  daltons. There are 9 ways for an rDNA set to form an out-of-register structure, and one way to form an in-register duplex. Assume that the rate of duplex formation is proportional to the number of complementary nucleotides. On this simple basis, the predicted ratio of out-of-register duplexes to in-register rDNA is  $9(1.5 \times 10^6 / 5 \times 10^7) = 0.27$ . We observe 0.5 out-of-register duplex sequences of length 5 kb per grid square (Table 2), so the observed ratios is  $0.5/7 \approx 0.07$ . Because of excluded volume effects, the simple assumption made about equal renaturation rate per complementary nucleotide for in-register and out-of-register duplexes is likely to overestimate the probability of out-of-register renaturation. We believe the fact that the observed number of out-of-register duplexes is about 1/4th of that calculated above is consistent with the hypothesis that they are rDNA sequences.

Linkage groups in rDNA. Grids made with self-renatured B. subtilis 168 (SPO2) DNA, with self-renatured B. subtilis ( $\phi 105$ ) DNA, and with heteroduplexes between lysogen and phage DNA's in all combinations were scanned. Approximately 20 singlets, 40 doublets and triplets, and eleven singlet and doublet structures adjacent to SPO2 prophage/phage heteroduplexes (4 of which were circular) were photographed and measured. (The linkage of rDNA sequences to prophage SPO2 is discussed later.) The mean length and standard deviation of the putative rDNA duplexes in this sample was  $4.83 \pm 0.25$  kb. This result is consistent with the interpretation that all the duplexes have the same length.

On the basis of the various out-of-register rDNA heteroduplex structures seen (Fig. 4), a set of four linkage groups, containing ten rDNA sets, is proposed, as shown in Fig. 4. The interpretation of each observed structure as a heteroduplex between two groups is indicated in the figure. All of the linkage groups are necessary to explain the observed structures, except that it is conceivable that linkage group D could contain A and/or B, or C in the spacer of length 55 kb. This is unlikely however. Suppose, for example, that the 55 kb spacer of D contained sequence B. One would expect, instead of structures j and k of Fig. 4, to see structures with a smaller single-strand loop, because of the well-known decrease in probability of ring closure with increasing ring size (Jacobson and Stockmayer, 1950). It is also possible that the linkage groups overlap. For example, as discussed later, we believe that rDNA sequences 5C<sub>1</sub> and 5D<sub>4</sub> are the same, so that groups D and C are combined in a still larger linkage group.

Linkage between rDNA and prophage SPO2. Density transfer and PBS-1 mediated transduction experiments show that the SPO2 prophage is closely linked to the erythromycin locus, *ery-1* (Inselburg et al., 1969). Density transfer and transformation experiments indicate that some rDNA genes are closely linked to the streptomycin and erythromycin resistance markers (Dubnau, Smith & Marmur, 1965), although the resolution of these experiments is much less than that of transduction experiments.

By careful scrutiny of grids prepared by hybridizing B. subtilis 168 (SPO2) lysogen DNA with S25 phage DNA and with a



mixture of R90 and wild-type SPO2 phage DNA's, we have seen structures with a prophage/phage heteroduplex linked to rDNA duplexes by single-strand DNA. The structures seen are depicted in Fig. 5. Plate V exhibits several electron micrographs. These observations show that an rDNA doublet which has a spacer of less than 600 base pairs is located 6.2 kb away from the prophage. The structure with a linear phage/prophage heteroduplex (Fig. 5, Plate V) shows that the S25 deletion loop is distal to the bacterial DNA:prophage DNA junction that is 6.2 kb from an rDNA gene. As shown in the figure, structures were also seen in which a single rDNA duplex is separated from the prophage by  $12.0 \pm 0.2$  kb. This structure can be explained if the rDNA sequence involved in duplex formation is the distal member of the doublet, whereas the proximal member, which is 6.2 kb away from the prophage, has remained single-stranded. The predicted spacing for this model is  $4.8 + 6.2 + (0.6) \approx 12$  kb.

Table 4 summarizes the measurements of the spacers in the several linkage groups. A notation to identify the different spacers and rDNA sequences is introduced in Fig. 4. Two very short spacers,  $sB_1$  and  $sD_1$ , are observed with apparent lengths of  $0.28 \pm .05$  and  $0.14 \pm .05$  kb. We do not believe that length measurements are sufficient to distinguish between them. A spacer of length 0.6 kb is also seen and is part of our proposed linkage group A. We believe that the 0.6 kb spacing can usually be distinguished from the shorter spacings, but there are some heteroduplexes in which the identification is not certain. Figure 4 shows two sets of hetero-

Fig. 5 Heteroduplexes between SPO2 prophage ( B. subtilis 168 (SPO2) ) and phage DNA's ( wild type, R90 and S25 ) adjacent to rDNA sequences.

Numbers are lengths in kilobases. a , b , c , d , and e were observed in lysogen / S25 heteroduplex preparations; d and f were observed in a lysogen / wild type, R90 heteroduplex preparation. A total of eleven molecules were seen, four of which had a circular prophage / phage DNA structure.

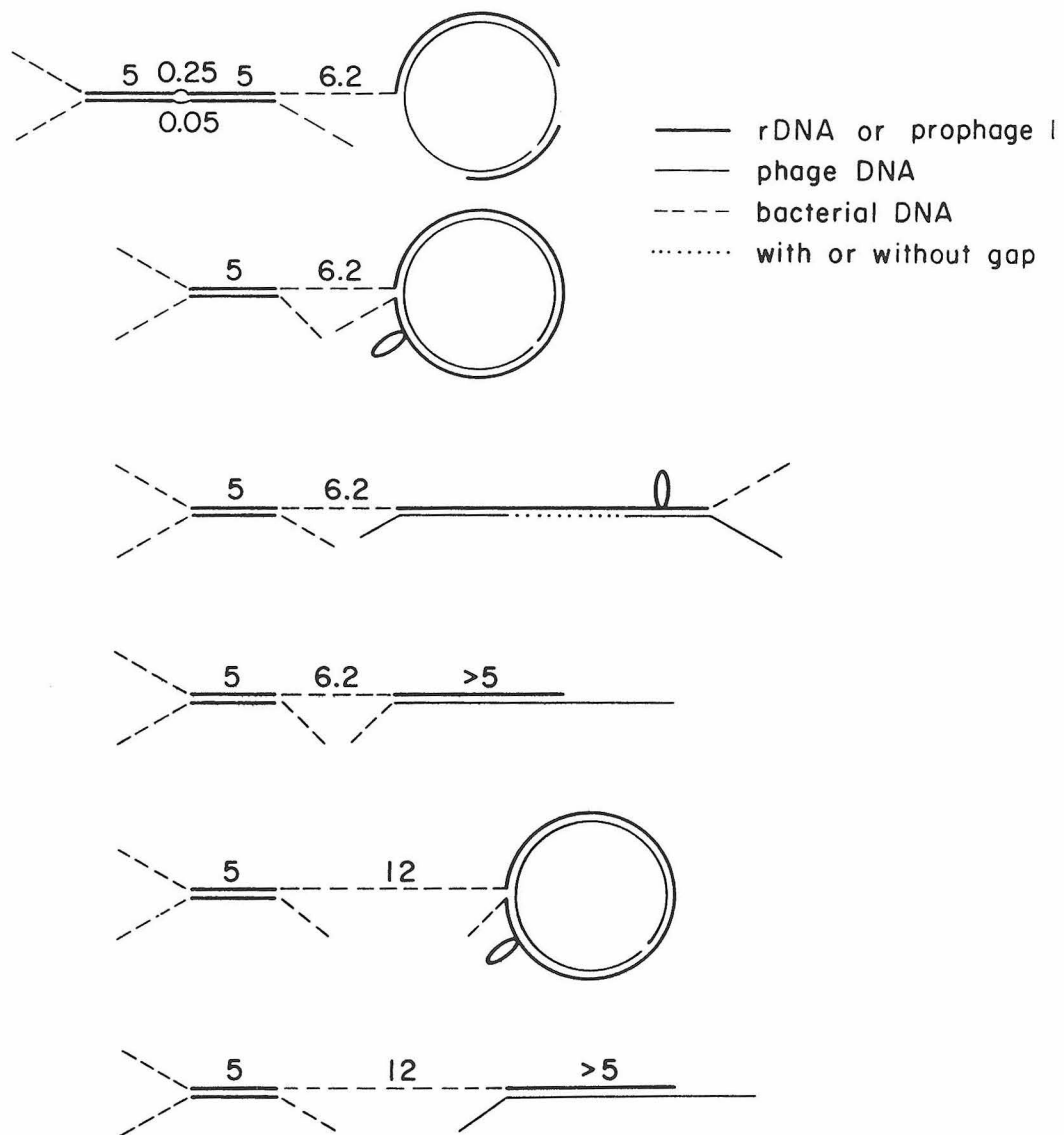


Fig. 5

Plate V Linkage of prophage SPO2 and ribosomal DNA.

- A Linear prophage / phage heteroduplex; separated from an rDNA singlet by a 6.2 kb single-strand spacer. The S25 deletion loop can be seen at the distal end of the rDNA, and is identified as  $\Delta$ . A schematic representation is shown in Fig. 5c.
- B Circular prophage/ phage heteroduplex; separated by 12.0 kb from an rDNA singlet. This is the molecule depicted in Fig. 5e. The S25 deletion loop is marked  $\Delta$ .
- C Circular prophage / phage heteroduplex; separated by 6.2 kb from an rDNA doublet. This is the molecule depicted in Fig. 5a. The inset tracing is an enlargement of the region around the very small spacer loop, with lengths 0.25 and 0.05 kb ( spacers sD<sub>1</sub> or sA<sub>1</sub>, and sB<sub>1</sub> or sD<sub>1</sub>, respectively ).

Arrows point to phage attachment site(s), rDNA, spacer DNA, and the junction of rDNA duplex and single-strand DNA. DNA's from B. subtilis 168 (SPO2) and phage SPO2 S25 were denatured and renatured as described in the text.

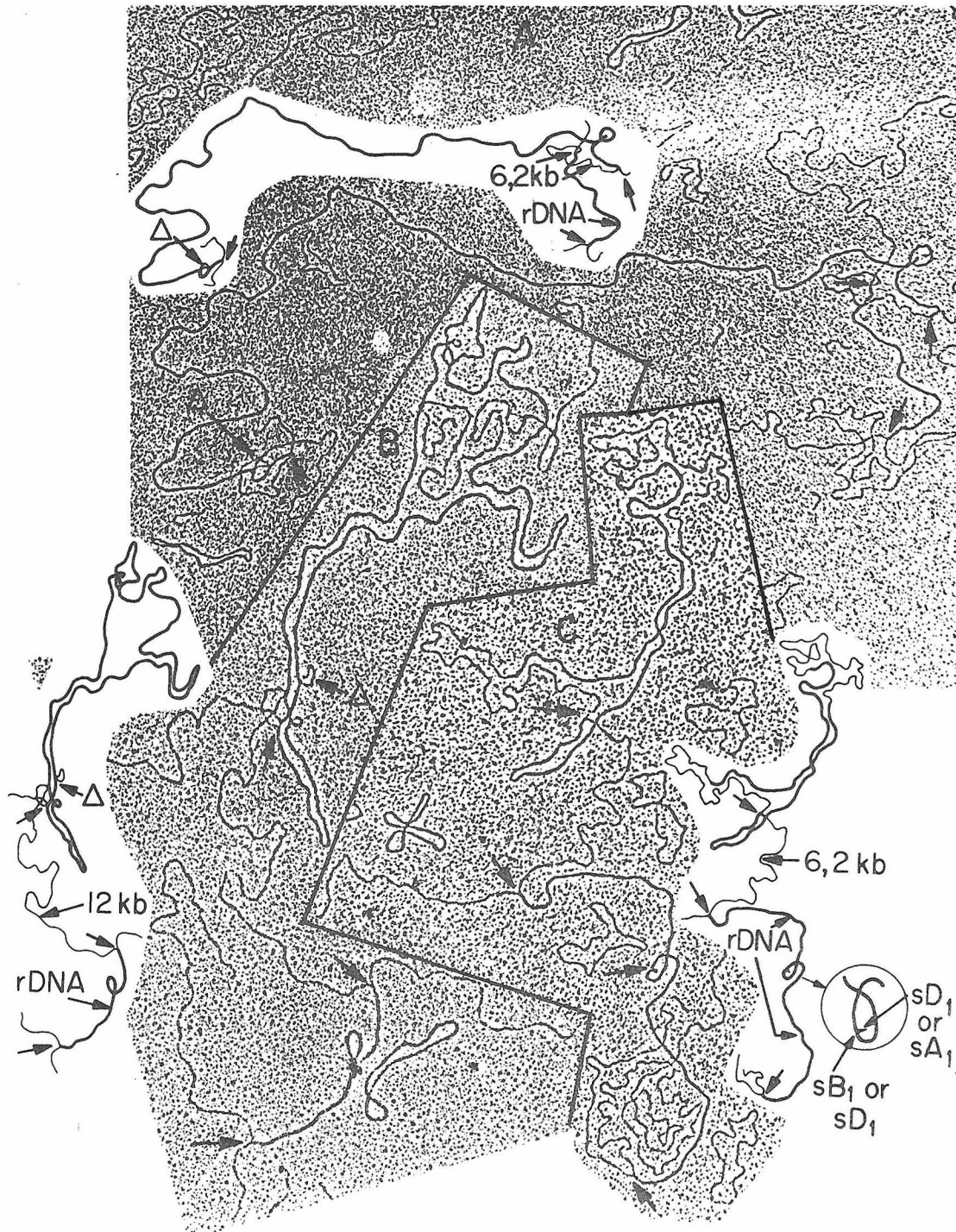


Plate V

TABLE 4

Lengths and standard deviations of ribosomal spacer DNA's\*

Spacer**	No. Averaged	Length (kb)	Standard Deviation (kb)
sA <sub>1</sub>	12	0.57	0.1
sB <sub>1</sub>	9	0.14	0.05
sC <sub>1</sub> (sD <sub>4</sub> )	17	14.97	0.68
sD <sub>1</sub>	6	0.28	0.05
sD <sub>2</sub>	11	1.04	0.15
sD <sub>3</sub>	7**	55.37	0.9
Prophage-rDNA	9	6.24	0.2
Prophage-rDNA	2	12.0	0.2

\* Single strand DNA connecting two  $4.83 \pm 0.25$  kb duplex segments as in Fig. 4.

\*\* Four were single stranded, (Fig. 4j, k, and m and Fig. 6f), three were partially duplex, partially single stranded (Fig. 6 a, b, c).

duplexes with two short spacers on opposite strands (4d and 4e). If the interpretation that heteroduplexes 4d and 4e are distinct is correct, groups A, B, and D must all exist. If 4e is really the same as 4d, then only A and D exist.

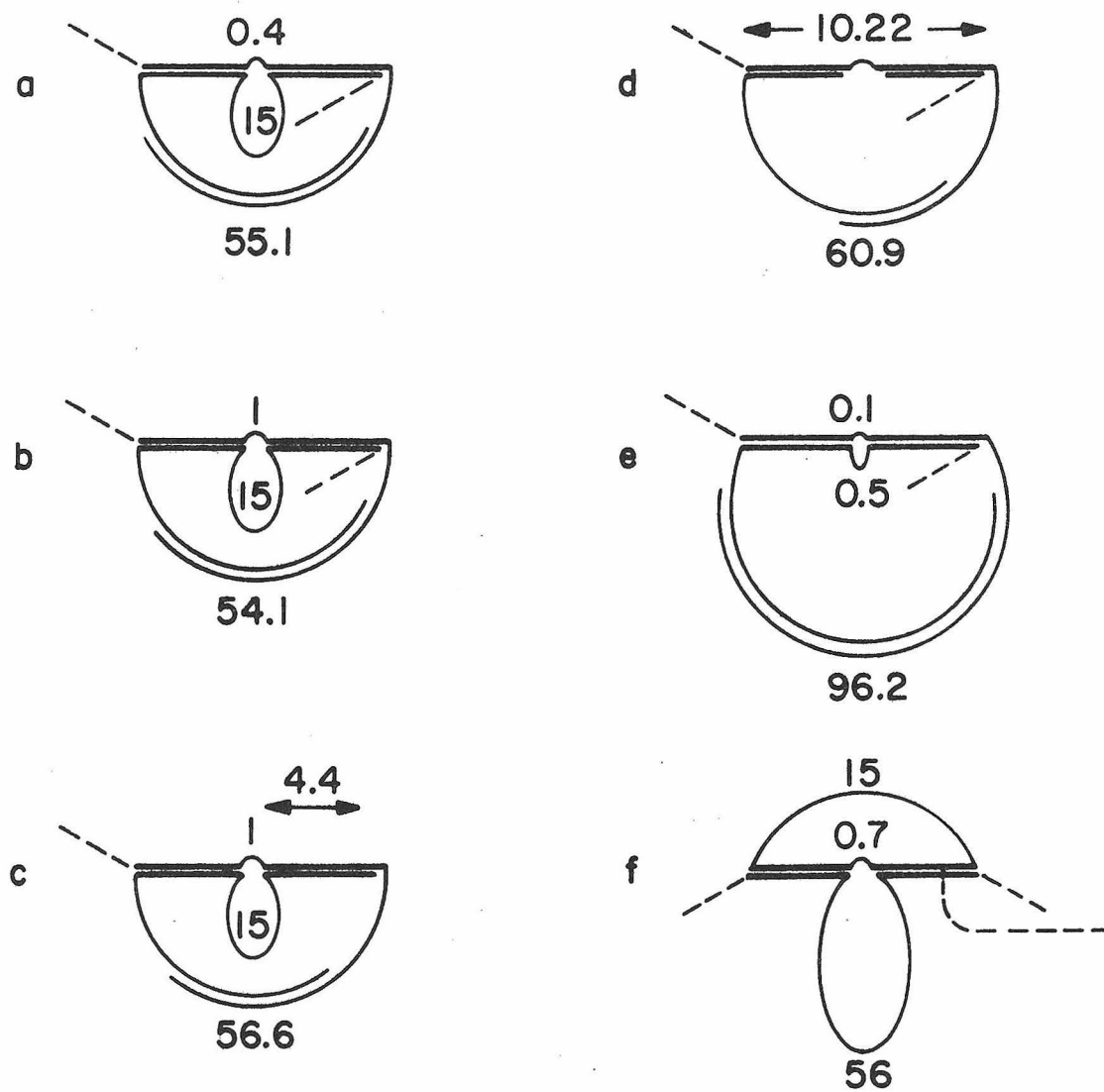
Most of the data reported in this and the preceding section were collected using lysogen DNAs prepared from cells grown directly from a slant. It is conceivable therefore that the several linkage groups with spacers of different length arose from heterogeneity in the cell populations of B. subtilis 168 (SPO2) and of B. subtilis 168 ( $\phi$ 105). A culture of cells was prepared from a single colony of B. subtilis 168 (SPO2). The DNA from this culture was self-annealed and hybridized to phage S25 DNA. Structures with spacers identical to those seen in Fig. 4 were seen as were structures with a 6.2 kb spacer separating rDNA and prophage. Therefore, we conclude that all the structures seen are due to sequence relations in a homogeneous B. subtilis chromosome.

Linking the linkage groups. Several circular structures which cannot be explained solely on the basis of the linkage groups of Fig. 4 are displayed in Fig. 6. Electron micrographs are shown in Plates VI, VII, and VIII. Structures b and c are the same. A substitution loop with two arms of length 15 kb and 1 kb, respectively, separates the two rDNA duplexes. Structure a seems to be different. One single-strand of the spacer loop is 15 kb, but the other is only 0.4 kb. For all three cases, four single-strand arms emerge from the ends. Two of these are joined in a circle of length about  $55 \pm 1$  kb. Part of this is duplex. We propose that the rDNA sequences 5C<sub>1</sub> and

Fig. 6 Observed circular structures with rDNA duplex regions.

The numbers are lengths in kb. The outside arms that are connected to form the circular structures are assigned arbitrarily to rDNA strands. Each structure was seen only once. a , b , e , and f are complete structures showing single strand arms at ends and substitution loops between the doublets. In molecule c , the right rDNA sequence is incomplete with a length of 4.4 kb only. In molecule d , since one of the rDNA strands is broken, only the end to end distance of the rDNA doublet is given.





— rDNA  
 — spacer DNA  
 ---- other bacterial DNA

Fig. 6

Plate VI   Circular out-of-register ribosomal DNA heteroduplex  
from B. subtilis 168 (SPO2).

This is the structure depicted in Fig. 6a and 7a or 7b.

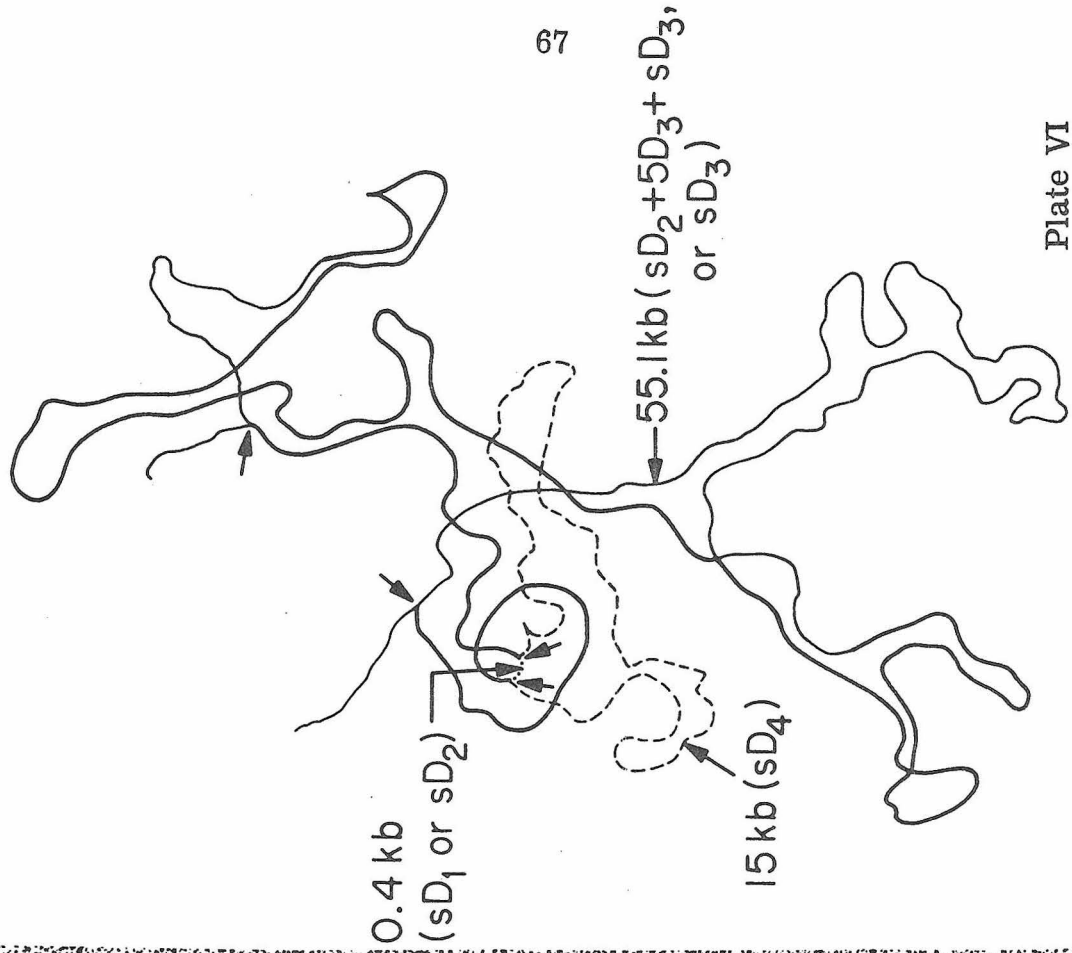
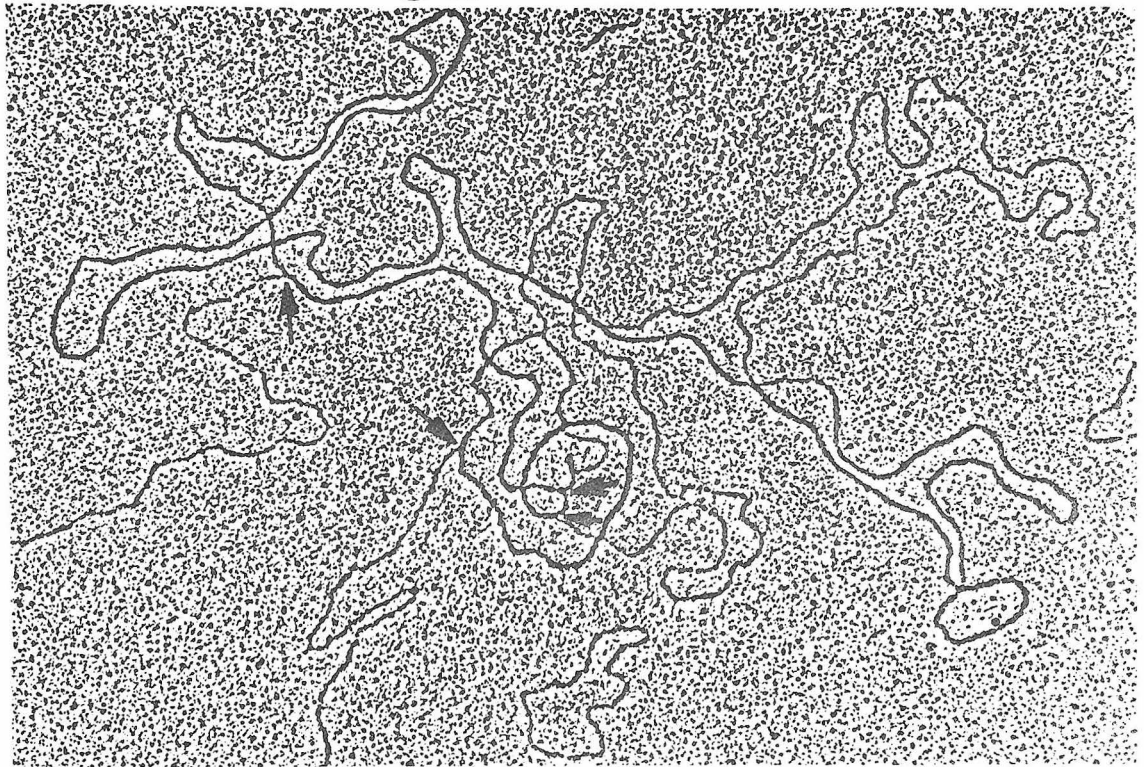


Plate VII   Circular out-of-register ribosomal DNA heteroduplex  
from B. subtilis 168 (SPO2).

This is the structure depicted in Fig. 6f and 7c<sub>1</sub> or 7c<sub>2</sub>.

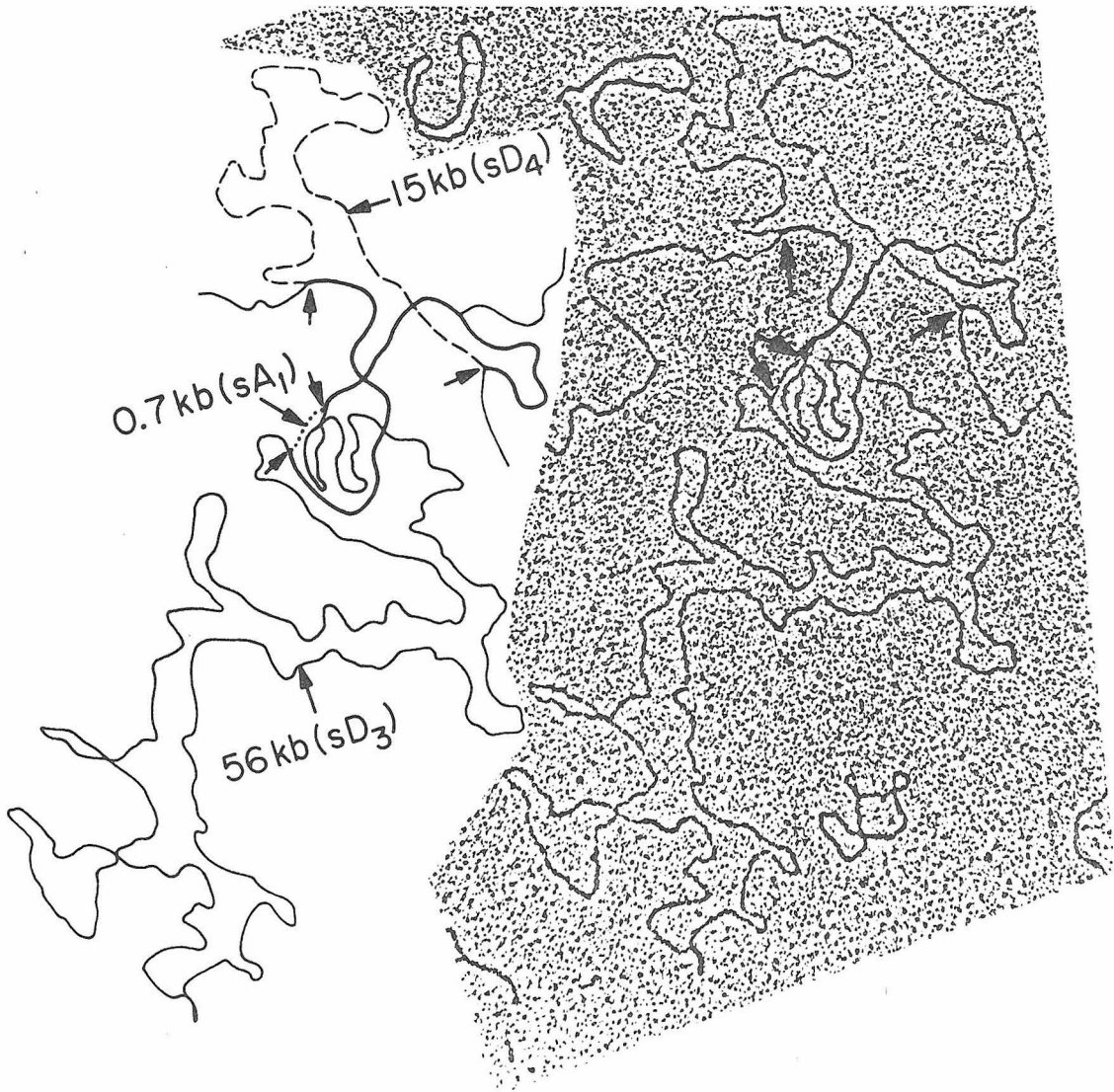


Plate VII

Plate VIII    Circular out-of-register ribosomal DNA heteroduplex  
from B. subtilis 168 (SPO2).

This is the structure depicted in Fig. 6e.

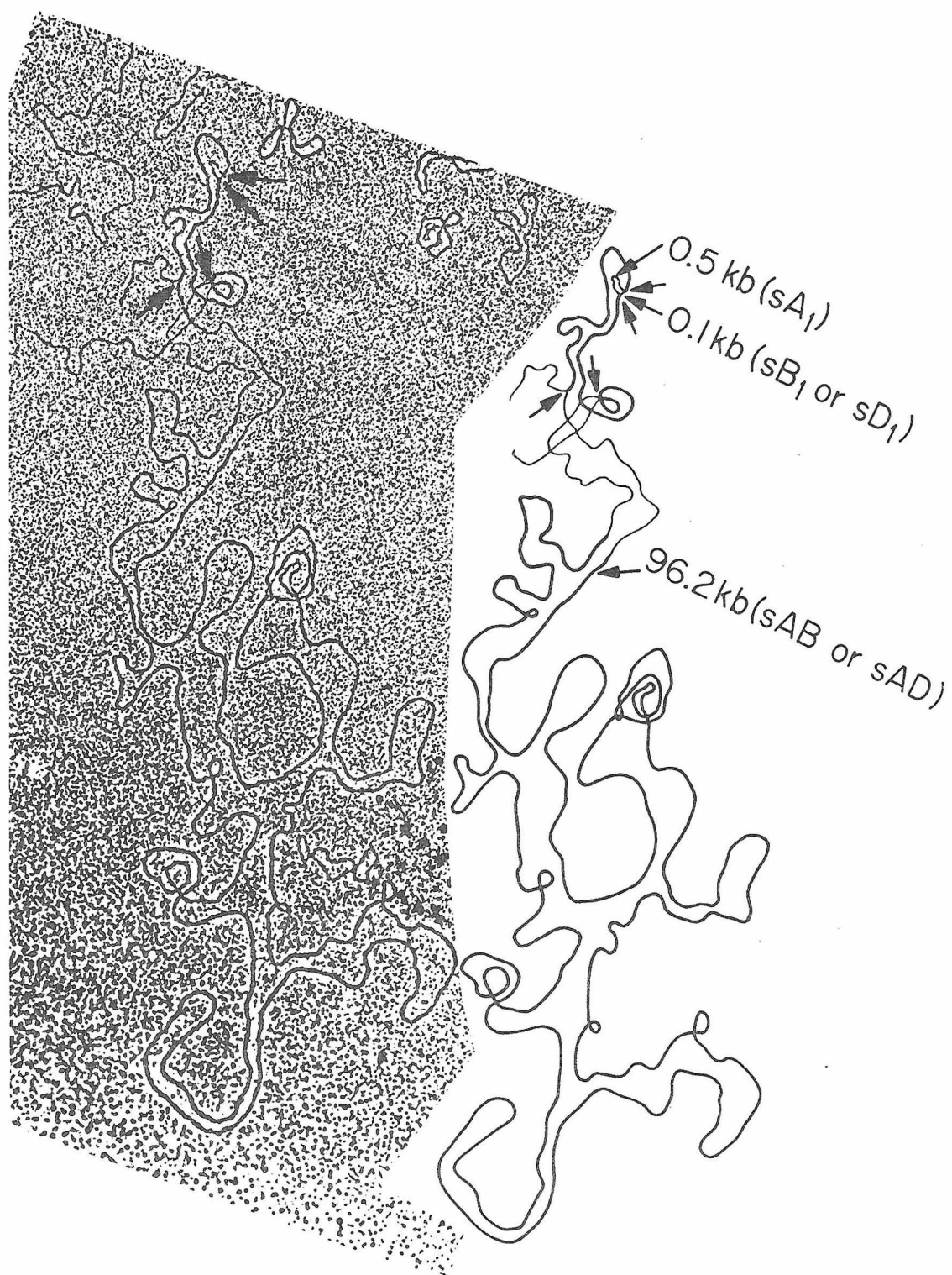


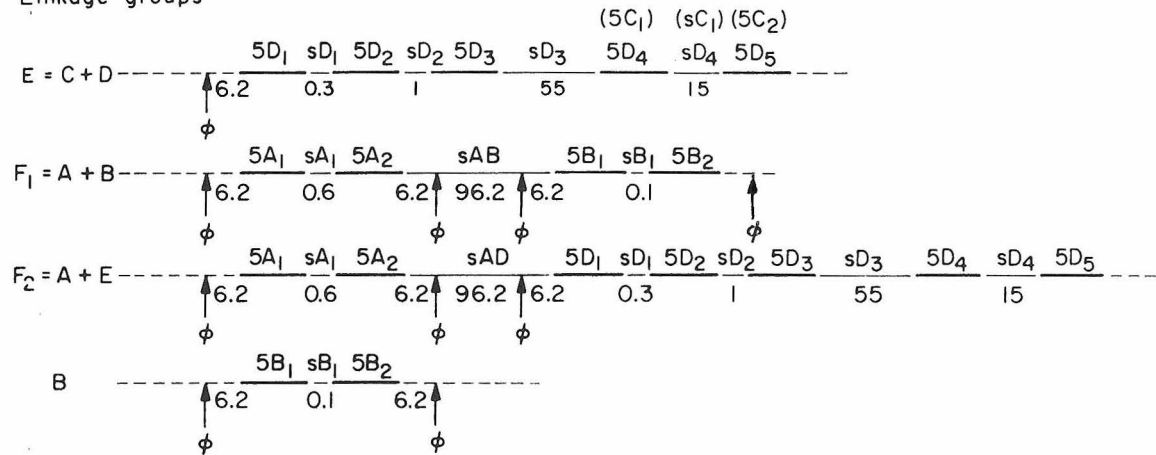
Plate VIII

Fig. 7 Larger rDNA linkage groups and circular heteroduplexes.

The larger linkage groups constructed by joining the groups of Fig. 4 are shown, as are possible positions for the SPO2 prophage. The composition of the larger groups in terms of the groups of Fig. 4 are shown to the left. The new spacer of length 96.2 kb formed by joining A and B or A and D is labeled sAB or sAD, respectively. Circular heteroduplex structures formed by renaturation of these linkage groups are shown below. The correlation between these structures and the observed structures of Fig. 6 is discussed in the text. A structure similar to a can be formed by an  $F_1/F_1$  or an  $F_2/F_2$  heteroduplex ( in the latter case, only the four left rDNA sequences are involved ) to give the observed structure 6 e . The several possible positions of prophage SPO2 are indicated by  $\phi$ . In each case, the junction between prophage DNA and bacterial DNA is located 6.2 kb away from the nearest rDNA sequences. The 1.4 kb DNA sequence deleted in the S25 mutant is situated distal to the rDNA sequences. Thus, the deleted segment is  $6.2 + 23.6 + (1.000 - 0.669) \times 38.6 = 42.6 \text{ kb to } 44.0 \text{ kb}$ , from an rDNA duplex ( see also Fig. 2 ).



## Linkage groups



## Structures

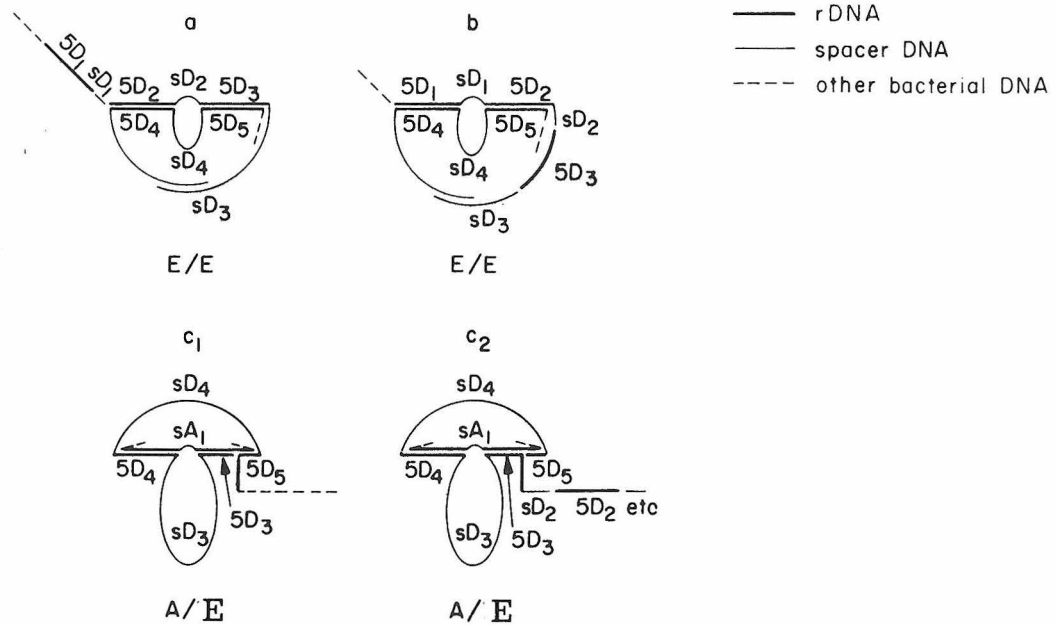


Fig. 7

5D<sub>4</sub> of Fig. 4 are the same, so that linkage groups D and C are part of a larger linkage group shown as E in Fig. 7. A way in which this combined linkage group can renature out-of-register with its own complement to give structures 6b and c is also shown in Fig. 7a.

Figure 7b shows one interpretation of the structure of Fig. 6a as another kind of E/E heteroduplex. In this interpretation, the single-strand spacer sD<sub>1</sub> of length 0.4 kb is one arm of the loop separating the two rDNA duplexes. The large circle, partly single-strand and partly duplex has a predicted length of 61 kb, which differs by two standard deviations from the experimental value of  $55.1 \pm 2.7$  kb. An alternative interpretation is that 6a is the same as 6b and 6c, with a large circular arm of 55 kb, and the small spacer is sD<sub>2</sub> of length 1 kb although it is measured as 0.4 kb. There is always some variability in heteroduplex structures, and particular features, especially single-strand segments, may be distorted in individual cases. A third possibility is that the sample of *B. subtilis* DNA is somewhat heterogeneous; in some chromosomes, sD<sub>2</sub> is of length 1 kb; in others, its length is 0.4 kb. Other explanations can be proposed. But since the structure has been seen only once, the various hypotheses cannot be tested with the data available. Structure 6f confirms the linkage of the 15 kb and 55 kb spacers as shown in Fig. 7c<sub>1</sub>, and c<sub>2</sub>.

Two other circular structures, each seen only once, are also shown in Fig. 6. 6d, with a partially duplex circular spacer of length 61 kb, is most likely formed by the E/E heteroduplex, 7b. However,

one of the long pairing strands is broken. The 15 kb spacer, sD<sub>4</sub>, is not seen. The total end-to-end length of the rDNA doublet is  $10.22 \pm 0.19$  kb. Since each rDNA duplex is  $4.83 \pm 0.25$  kb, the calculated length of the spacer is  $0.6 \pm 0.5$ , which is consistent with its assignment as sD<sub>1</sub>. The structure of Fig. 6 e has a large circular partially duplex spacer of length 96.2 kb. Its explanation requires a new linkage group, either 7 F1 or 7 F2, in which linkage group A is connected by a 96 kb spacer to either B or D. Since structure 6e has been seen only once, the evidence for either linkage group 7 F1 or F2 is not compelling. However, the micrograph of the heteroduplex, 6e (Plate VIII), shows a good, clear structure. We are inclined to believe that this structure is real and that one of the linkage groups 7F is real. We have already pointed out that linkage group B may be a part of D, with sB<sub>1</sub> and sD<sub>1</sub> identical. In that case, 7 F1 is part of 7 F2.

The SPO2 prophage sequence, of length 39 kb, is 6.2 kb from one of the rDNA doublets containing a short spacer (sA<sub>1</sub>, sB<sub>1</sub>, or sD<sub>1</sub>) (Plate Vc). Several possibilities for its position are shown in Fig. 7. It could be to the left of the rDNA sequence 5D<sub>1</sub> in linkage group 7 E. It could be to the left of 5A<sub>1</sub> or 5B<sub>1</sub>, or to the right of 5A<sub>2</sub> or 5B<sub>2</sub>, in any of the linkage groups proposed, except that if 5B<sub>2</sub> and 5D<sub>2</sub> are identical, the prophage could not be 6.2 kb to the right of this rDNA set.

Absence of inverted repeats. A sample of *B. subtilis* 168 (SPO2) DNA in CsCl was denatured by alkali, reneutralized, and dialyzed into 60% formamide, 0.1 M Tris, 0.01 M EDTA, pH 8.5 at 4°C, and then into 0.01 M Tris, 0.001 M EDTA, pH 8.0 at 4°C, as described in Materials and Methods. This solution was diluted into the spreading solution for electron microscopy. Previous experience in this laboratory has shown that, under these circumstances, inverted repeat sequences on a single strand of DNA "snap-back" to form duplex, whereas a sample without inverted repeats contains only a few per cent of long duplex DNA due to rejoining of complementary strands or conceivably to duplex that did not denature. An inverted repeat structure is recognizable because there is a single-strand circle at one end of the duplex. Examples of such structures are given by Sharp, Hsu, Ohtsubo, and Davidson (1972) and by Sharp, Cohen, and Davidson (1972).

No structures due to inverted repeats were seen in the *B. subtilis* DNA, after treatment as described above. The average single-strand length in the sample was about 130 kb. We conclude that any two rDNA cistrons on the *B. subtilis* chromosome that are within this distance of each other have the same polarity.

Comments on the kinetics of renaturation. In Table 3, it is reported that for *B. subtilis* DNA at a concentration of 10  $\mu\text{g/ml}$  ( $3.0 \times 10^{-5}$  M in nucleotide units), 25% renaturation occurs in two hours.

The standard second-order renaturation equation for DNA may be

written as

$$\underline{R}/(1 - \underline{R}) = \underline{k}_2 \underline{c}_0 t / 2 \quad (1)$$

where  $\underline{R}$  is the fraction renatured,  $\underline{k}_2$  is the second-order renaturation rate constant, and  $\underline{c}_0$  is the initial concentration of denatured DNA in nucleotide units (see Eq. 9 of Wetmur & Davidson, 1968). We thus calculate that for B. subtilis DNA in the 70% formamide, 0.25 M NaCl, medium used for renaturation at ca. 25°C,  $\underline{k}_2 = 3.1 \text{ M}^{-1} \text{ sec}^{-1}$ . This is for a DNA of complexity,  $\underline{N} = 5.9 \times 10^6$  nucleotide pairs ( $3.9 \times 10^9$  daltons), and single-strand length,  $\underline{L}$ , of ca.  $1.3 \times 10^5$  nucleotides. From the equation of Wetmur & Davidson (1968) for an aqueous medium with  $[\text{Na}^+] = 0.25 \text{ M}$  at  $\underline{T}_m - 25^\circ$

$$\underline{k}_2 = 7.5 \times 10^4 \underline{L}^{0.5} / \underline{N} \quad (2)$$

we calculate a predicted  $\underline{k}_2$  of  $4.5 \text{ M}^{-1} \text{ sec}^{-1}$ . The observed rate for B. subtilis DNA is therefore consistent with previous observations.

The above discussion suggests that

$$\underline{k}_2 = 5 \times 10^4 \underline{L}^{0.5} / \underline{N} \quad (3)$$

under the renaturation conditions used. For SPO2 DNA ( $\underline{N} = \underline{L} = 4.0 \times 10^4$  nucleotides), Eq. (3) gives  $\underline{k}_2 = 250 \text{ M}^{-1} \text{ sec}^{-1}$ .

The lysogen DNA, phage DNA heteroduplex renaturations were conducted at respective DNA concentrations of 10  $\mu\text{g/ml}$  and 0.5  $\mu\text{g/ml}$ . The calculated concentration of prophage DNA (assuming one prophage per bacterial chromosome) is 0.07  $\mu\text{g/ml}$ , thus phage DNA

is in excess. At the phage DNA concentration used, Eq. 1 predicts  $\underline{R} = 0.6$  for the degree of self-renaturation of the phage DNA in 2 hours; we thus estimate an average phage DNA concentration,  $\langle \underline{c}_\phi \rangle$ , of  $0.3 \mu\text{g/ml}$  or  $1.0 \times 10^{-6} \text{ M}$  during the renaturation experiment.

Assume  $k_2 = 250 \text{ M}^{-1} \text{ sec}^{-1}$  for renaturation of prophage DNA with phage DNA. For the formation of prophage/phage heteroduplexes we may then write

$$-\frac{dc_{p\phi}}{dt} = k_2 c_{p\phi} \underline{c}_\phi \approx k_2 c_{p\phi} \langle \underline{c}_\phi \rangle \quad (4)$$

where  $\underline{c}_{p\phi}$  and  $\underline{c}_\phi$  are concentrations of single-strand prophage DNA and phage DNA, respectively. The equation states that the disappearance of prophage DNA is approximately first-order, and with our estimate of  $\langle \underline{c}_\phi \rangle$  of  $1.0 \times 10^{-6} \text{ M}$ , we calculate that 84% of the prophage DNA should have formed heteroduplexes with the phage DNA. With an estimate of 2.75 bacterial chromosomes per electron microscope grid square, one would therefore expect to find about 4 heteroduplexes per grid square. Our observed frequency was about 0.7 (Table 2). This discrepancy may indicate that many of the heteroduplex structures were either broken or so tangled or overlapped by other DNA that they were difficult to recognize, or may be due to defects in the kinetic analysis above. Nevertheless, the results of the kinetic analysis may be useful as guides in other similar experiments.

#### 4. Discussion

The results reported here show quite conclusively by a physical method that the prophage of SPO2 is a circular permutation of the phage; the att site is mapped at a position 61.2% from one end of the mature phage DNA. Rather convincing negative evidence that  $\phi$ 105 integrates without detectable circular permutation has been offered. This conclusion agrees with genetic evidence that the marker order in phage and prophage  $\phi$ 105 are colinear (Armentrout & Rutberg, 1970). Only a limited number of markers are available for mapping in this study; the possibility remains that the att site is outside these markers but not at the ends. Based on the  $\lambda$  precedent, duplexes with complementary single-strand ends of length greater than 20 would have a very high probability of circularizing (Wang & Davidson, 1968). From our physical studies, it is probable therefore that the att site is within 20 nucleotides of the ends of  $\phi$ 105.

We have observed a low frequency of duplex structures in self-renatured B. subtilis 168 (SPO2) DNA and in B. subtilis 168 ( $\phi$ 105) DNA. The uniform length of the duplex regions of  $4.83 \pm 0.25$  kb and the disappearance of these structures when the DNA is renatured in the presence of B. subtilis rRNA supports the conclusion that these duplexes are due to the rDNA sequences in the bacterial chromosome.

The accuracy and the certainty of the inferences about the linkage groups of the rDNA are not as high as desired, because of

the low frequency of the occurrence of the structures. We believe that the data strongly support the linkage groups shown in Fig. 4 and the larger linkage group E of Fig. 7. Individual heteroduplexes have been seen which suggest other linkage groups (7 F1 or 7 F2). One structure on an electron microscope grid could be due to accidental overlaps or aggregation of strands or to other artifacts; thus the evidence for such other linkage groups is not compelling.

Hybridization data suggest 9 to 10 rDNA sets per B. subtilis chromosome (Smith et al., 1968). If linkage groups E and F1 exist, and are independent, they contain nine rDNA sets. If they are joined as in F2, only 7 rDNA sets are included. If as discussed previously, B is not part of D, it accounts for two more rDNA sets, to give a total of 9. Other rDNA sets may exist as unlinked singlets. The results do not exclude the possibility of several additional distinct linkage groups of type A or of type B, provided that they are not joined in the larger linkage groups of Fig. 7.

The length distribution of the rDNA duplexes is rather homogeneous; the mean length is 4.83 kb and the standard deviation is 0.25 kb. As previously mentioned, the expected length of a 16S plus a 23S rRNA sequence is about 4.8 kb. Since we saw no duplexes of length 3.2 kb or 1.6 kb we conclude that each 16S rDNA sequence is linked to a 23S rDNA sequence, and vice versa. If there are spacer sequences between the 16S and 23S sequences of a set, they must be homologous for the different rDNA sets or they must be less than 50 nucleotides in length. Otherwise, a small substitution loop



would have been observed in the rDNA duplexes. Furthermore, the order of 16S and 23S sequences must be the same in all the rDNA sets.

The observed total length of the duplexes (4.8 kb) and the combined molecular lengths of 16S and 23S rDNA (4.8 kb) are very close. There are experimental errors in both numbers. Nevertheless, the data suggest that the total amount of homologous spacer sequences (and possibly 5S DNA sequences, see below) in a 16S and 23S rDNA set must be less than about 750 nucleotides (three standard deviations).

Membrane filter hybridization experiments indicate that there are only 3 to 4 5S RNA genes per B. subtilis chromosome, whereas there are 9 to 10 16S and 23S genes (Smith et al., 1968). The 5S DNA sequences are more closely linked to 23S than to 16S sequences (Colli et al., 1971).

It is pertinent to discuss the question of whether in B. subtilis there really are fewer 5S rRNA genes than there are 16S and 23S rRNA genes. In E. coli the 16S-23S-5S rDNA sequences constitute a transcription unit (Dolittle & Pace, 1971). There is one molecule each of 5S, 16S, and 23S rRNA's in the 70S ribosome (Kurland, 1960; Rosset & Monier, 1963; Rosset, Monier & Julian, 1964). There are about equal numbers of the three RNA's in the cell and the same number of molecules are synthesized in a given time in a cell (Galibert, Lelong, Larsen & Boiron, 1967). Pace & Pace (1971) report that, in E. coli, there are as many 5S rDNA genes as there

are 16S and 23S, whereas Attardi, Huang, & Kabat (1965) report that there are about twice as many 5S rDNA sequences present. Thus the data are somewhat conflicting. But since one RNA molecule of each type occurs in a ribosome and since the three RNA molecules seem to be coordinately transcribed, the simplest situation would be that there is the same number of genes for 5S RNA as for 16S and 23S. The same argument applies by analogy to the *B. subtilis* ribosomal RNA species. If the number of 5S RNA genes is less than the number of 16S and 23S, but the numbers of stable molecules in the cell are the same, there must be some mechanism by which the synthesis rate, per gene, is greater for the 5S species.

Do our physical data bear on this question? 5S RNA has a length of about 120 nucleotides. All of our observed duplexes have the same length within  $\pm 750$  nucleotides at a confidence level of 0.997. Thus, it would be possible for some 16S and 23S rDNA sets to contain a 5S rDNA, whereas others do not, and still give heteroduplexes that fit the observed length distribution, provided the length of the homogeneous spacer between 23S rDNA and 5S rDNA is less than about 600 nucleotides. Furthermore, if some 16S and 23S rDNA sets contain a 5S sequence and others do not, the 5S sequence could not lie between the 16S and 23S sequences in the set, otherwise an insertion loop would be seen in some of the rDNA duplexes. Thus, if the number of 5S sequences is less than the number of 16S plus 23S sequences, the order in a linked group must be 16S-23S-5S in agreement with the conclusion of Colli *et al.* (1971).

Previous mapping experiments show that the rRNA genes are in the general region of a number of antibiotic resistance markers (Smith et al., 1968). Some or all of the latter are probably on genes for the synthesis of ribosomal proteins. Thus, there is evidence for some clustering of ribosomal genes in B. subtilis. The observation of a cotransduction frequency by phage PBS-1 of a marker on prophage SPO2 and of ery-1 of 60 to 65%, suggests that these markers are separated by about 40 to 35% of the PBS-1 chromosome or ca. 110 kb (Hunter, Yamagishi & Takahashi, 1967). Our physical observation that a doublet is very close (6.2 kb) to the SPO2 prophage thus adds credibility to our overall interpretations.

The most direct way to map ribosomal genes would be to observe DNA:rRNA duplex regions along a single DNA strand--preferably with a prophage SPO2/phage duplex region as a reference point. The difference in appearance between DNA/RNA hybrids and single-strand DNA is less than that between DNA/DNA duplexes and single-strand DNA (Davis & Hyman, 1970; Wu, Davidson, Aloni & Attardi, 1972). In the present experiments, RNA:DNA hybrid regions are such a small fraction of the DNA on the grid that we could not recognize them.

In all electron microscope heteroduplex work, the operator scans grids and selects structures to be photographed because they appear to be significant and because they are sufficiently untangled so that they can be traced and interpreted.. The distinction between artifacts and real structure is based on the frequent occurrence of

features with a certain topology and on the reproducibility of length measurements of these features.

In general, in studying complicated structures which occur at a low frequency there is a danger of arriving at incorrect conclusions because of the subjective selection of molecules to be photographed and an element of subjective interpretation in tracing the molecules. We find it gratifying however that, by persistent scanning of samples, it is possible to reach conclusions, such as those presented here, with a high confidence level.

ACKNOWLEDGMENT

We are indebted to Dr. LuBelle Boice for bacterial stocks, advice, and encouragement. We wish to especially acknowledge our indebtedness to Mr. Rufus Skillern of UCLA who generously provided us with samples of the deletion mutant phages S25 and R90, even though he had not completed his own studies and characterization of these phages and DNA molecules. Without his DNA samples, much of our work would have been impossible. This research has been supported by grant GM 10991 from the National Institutes of Health.

## REFERENCES

- Armentrout, R. W. & Rutberg, L. (1970). J. Virol., 6, 760.
- Attardi, P. C., Huang, P. C. & S. Kabat, S. (1965). Proc. Nat. Acad. Sci. U. S., 53, 1490.
- Chow, L. T., Boice, L. & Davidson, N. (1972). J. Mol. Biol., in press.
- Colli, W. & Oishi, M. (1969). Proc. Nat. Acad. Sci. U. S., 64, 642.
- Colli, W., Smith, I. & Oishi, M. (1971). J. Mol. Biol., 56, 117.
- Dabnau, D., Smith, I. & Marmur, J. (1965). Proc. Nat. Acad. Sci. U. S., 54, 724.
- Davis, R. W. & Hyman, R. W. (1970). Cold Spring Harbor Symp. (1970), 35, 269.
- Dolittle, W. F. & Pace, N. R. (1971). Proc. Nat. Acad. Sci. U. S., 68, 1786.
- Eberle, H. & Lark, K. G. (1967). Proc. Nat. Acad. Sci. U. S., 57, 95.
- Galibert, F., Lelong, J. C., Larsen, Ch. J. & Boiron, M. (1967). Biochim. Biophys. Acta, 142, 89.
- Hsu, M. T. & Davidson, N. (1972). Proc. Nat. Acad. Sci. U. S., in press.
- Hunter, B. I., Yamagishi, H. & Takahashi. (1967). J. Virol., 1, 841.
- Inselburg, J. W., Eremenko-Volpe, T., Greenwald, L., Meadow, W. L. & Marmur, J. (1969). J. Virol., 3, 627.

- Jacobson, H. & Stockmayer, W. H. (1950). J. Chem. Phys., 18, 1600.
- Kurland, C. G. (1960). J. Mol. Biol., 2, 83.
- Oishi, M. & Sueoka, N. (1965). Proc. Nat. Acad. Sci. U. S., 54, 483.
- Oishi, M., Oishi, A. & Sueoka, N. (1966). Proc. Nat. Acad. Sci. U. S., 55, 1095.
- Pace, B. & Pace, N. R. (1971). J. Bact., 105, 142.
- Rosset, R. & Monier, R. (1963). Biochim. Biophys. Acta, 68, 653.
- Rosset, R., Monier, R. & Julian, J. (1964). Bull. Soc. Chim. Biol., 46, 87.
- Rutberg, L. (1969). J. Virol., 3, 38.
- Sharp, P. A., Hsu, M. T. & Davidson, N. (1972). J. Mol. Biol., in press.
- Sharp, P. A., Hsu, M. T., Ohtsubo, E. & Davidson, N. (1972). J. Mol. Biol., in press.
- Sharp, P. A., Cohen, S. & Davidson, N. (1973). J. Mol. Biol., to be published.
- Smith, I., Dubnau, D., Morell, P. & Marmur, J. (1968). J. Mol. Biol., 33, 123.
- Wang, J. C. & Davidson, N. (1968). Cold Spring Harbor Symp., 33, 409.
- Wetmur, J. G. & Davidson, N. (1968). J. Mol. Biol., 31, 349.
- Wu, M., Davidson, N., Attardi, G. & Aloni, Y. (1972). J. Mol. Biol., in press.

### Part III

#### Electron Microscope Mapping of $\lambda$ dv DNA's



The genetic sequences of seventeen different  $\lambda$ dv's obtained from Dr. D. E. Berg have been mapped on  $\lambda$  DNA by the electron microscope heteroduplex method. The physical results are in agreement with the genetic analyses of the contents of the  $\lambda$ dv's. Different  $\lambda$ dv's exist predominantly either as monomers, or as dimers, or as trimers in recA<sup>-</sup> carrier cells. Higher order oligomers are also found in all  $\lambda$ dv preparations. Most of the  $\lambda$ dv's originated from  $\lambda$  phage carrying the nin deletion are observed to contain either an inverted complete duplication (and are called inverted dimers) or to contain inverted partial duplications (and are called amphimers). That is, the  $\lambda$ dv's carry a duplication of all or some of the sequences in an inverted order on the same strand. There are four types of inverted repeat structures: (1) complete inversion and duplication, (2) partial inversion and duplication with a long (5.7% of  $\lambda^+$ ) non-inverted, non-duplicated (or unique) sequence on the left end covering the immunity region (oriented according to the  $\lambda$  map), (3) partial inversion and duplication with a long (5.1%) and a short (2.2%) unique sequence on the left and right ends respectively, (4) partial inversion and duplication with a very short (<0.7%) unique sequence on the right end of the  $\lambda$ dv's near the nin deletion. When the closed circular molecules of these inverted  $\lambda$ dv's are nicked lightly, denatured by alkali, and then reneutralized, the complementary sequences on the same DNA strand "snap back" to form double-stranded linear molecules. The unique sequences, if any, appear as single-stranded loops at one or both ends when the above treated DNA preparations are mounted by the formamide technique. Lambda phages that do not carry the nin deletion do not

produce  $\lambda$ dv's with inverted repeats. Denaturation of the open circular molecules of these non-inverted  $\lambda$ dv's generates, as expected, only single-stranded circular and single-stranded linear molecules when mounted by the formamide technique. The nin deletion is believed to have caused somehow the formation of the  $\lambda$ dv's containing inverted repeats. The left end points of many  $\lambda$ dv's map around 73% on the  $\lambda$  map while the right boundaries of many  $\lambda$ dv's map in the vicinity of the nin deletion which extends from 83.8 to 89.6% on  $\lambda^+$ , indicative of two regions of high recombination frequency.

$\lambda$ dv (for  $\lambda$  phage defective virulent) plasmids were first discovered by Matsubara and Kaiser (1968) by repeated infection of E. coli with virulent mutants of  $\lambda$  phage. Dr. Douglas E. Berg has, since then, isolated many more  $\lambda$ dv's by infecting E. coli cells with various  $\lambda v^+$  or  $\lambda vir$  mutant phages with or without prior UV irradiation of the phage particles (Berg and Kaiser, 1972; Berg, personal communication). The plasmids exist autonomously with about 50 copies per carrier cell.  $\lambda$ dv carriers can be isolated because of their immunity to superinfecting  $\lambda$  and  $\lambda vir$  phages. However,  $\lambda$ dv from  $\lambda v^+$  phage is not immune to  $\lambda vir$ . Genetic tests indicate that all  $\lambda$ dv's have deleted 80% or more of the  $\lambda$  genome. They always contain the  $\lambda$  replication genes O and P and the origin of replication (Schnös and Inman, 1970; Stevens, Adhya, and Szybalski, 1971).

The mechanism of  $\lambda$ dv formation is not clear. Berg and Kaiser (1972) have postulated that they are formed by illegitimate recombination (Franklin, 1971) during  $\lambda$  DNA replication. The present electron microscope mapping study was undertaken in order to compare the physical map with the genetic map, and to look for structures not readily recognized by genetic analyses. In particular, it was hoped that precise information about the left and the right end points of the  $\lambda$ dv's might provide insight into the recombination events leading to their formation.

### Experimental Procedures

1. Bacterial strains for growing  $\lambda$  phages are listed in Table 1.
2. Lambda phages used are presented in Table 2.
3. All  $\lambda$ dv carrying strains, as well as the available information about the genetic contents of the  $\lambda$ dv's provided by Dr. D. E. Berg, are given in Table 3.
4. Preparations of bacteriophages.

#### A. Preparation of $\lambda$ phage grown in broth.

##### (1) Single-plaque purification of old phage stock.

Aliquots of phage stock were diluted in TMG (0.01 M Tris, 0.01 M  $\text{MgSO}_4$ , pH 7.4, 0.1% gelatin W/V) and adsorbed to E. coli C600 at room temperature and then plated on agar plates at 37° C. Top agar (3 ml/plate) contained 1% Bacto-trypton (Difco), 0.65% agar and 0.5% NaCl. Bottom agar contained 1% trypton, 1% agar and 0.5% NaCl. One clear plaque was picked with a sterilized inoculating needle and the phages were suspended in TMG for at least two hours. Chloroform could be added to prevent bacterial contamination.

##### (2) Preparation of new phage stock.

The phage suspension from a single plaque was titered and then plated on several agar plates to give about 2000 plaques/plate. The plates were incubated in a 37° C oven until the plaques were about to become confluent. One ml of chloroform per plate was then added

TABLE 1

Bacterial Strains for Growing  $\lambda$  Phages

<u>E. coli</u> Strains	Remarks
C600	for plating on agar
K12 W3110	for growing in broth
829S	for plating duplicate phage on agar; a derivative of W3101 carrying the WGS6 episome (Spiegelman, 1971); derepressed for $\lambda$ <u>N</u> gene when grown at 42° C for 3 generations. Compound phages make clear plaques on it, whereas segregants make turbid ones (Berg, 1971).

TABLE 2

Lambda Phages

Phages	Source	Method of Preparation
<u>c</u> <sub>26</sub>	This lab	B
<u>b</u> <sub>2</sub> <u>b</u> <sub>5</sub> <u>c</u>	This lab	A
<u>b</u> <sub>538</sub> <u>i</u> <sup>21</sup>	D. E. Berg	B
p <u>gal</u> 8 <u>vir</u> <u>cII</u> <u>nin</u>	D. E. Berg	B
p <u>gal</u> 8 <u>cI</u> sus 34	D. E. Berg	B
<u>b</u> <sub>221</sub> <u>cI</u> 857 p g q 4 *	D. E. Berg	B
Compound (duplicate)		
phages		
<u>b</u> <sub>538</sub> <u>i</u> <sup>21</sup> sus <u>P</u> -dv 1	D. E. Berg	C
<u>b</u> <sub>538</sub> <u>i</u> <sup>21</sup> sus <u>P</u> -dv 249	D. E. Berg	C
<u>b</u> <sub>538</sub> <u>i</u> <sup>21</sup> sus <u>P</u> -dv 292	D. E. Berg	C

\* Parental phage of  $\lambda$ dv 120. It contains gal genes, but is extremely unstable; over 95% of phage particles have lost the gal genes.

TABLE 3

 $\lambda$ dv's and their Genetic Information

Carrying Strains	Genetic End Points*	Right	Physical Size**, % of Deletion Required for Compound Phage Formation	Parent Phage
KM 424 <sup>s</sup> (dv 1)	73 <sup>†</sup>	87.5 <sup>†</sup>	≈	(Matsubara & Kaiser, 1968)
866 <sup>s</sup> (dv 21)			≈	( $\lambda$ vir <u>nin</u> <sup>-</sup> ) <sup>Δ</sup>
866 (dv 120) <sup>‡</sup>			21%	Fig. 10
866 (dv 154)	left of <u>i<sub>434</sub></u>	right of <u>nin</u> deletion	≈	$\lambda$ p gal 8 vir <u>cII</u> <u>nin</u>
866 (dv 203)			11%	$\lambda$ gal 8 <u>cI</u> 857 <u>nin</u> at 42° C
866 (dv 266)	left of <u>i<sub>434</sub></u> <sup>ΔΔ</sup>	right of <u>nin</u> deletion	≈, 5%	$\lambda$ p gal 8 vir <u>cII</u> <u>nin</u>
866 (dv 104)	between the left bound- aries of <u>i<sub>21</sub></u> and <u>i<sub>434</sub></u>	right of <u>Q117</u> left of <u>S<sub>7</sub></u>	>, 21% <sup>++</sup>	$\lambda$ p gal 8 vir <u>cII</u> <u>nin</u>

TABLE 3 (continued)

866 (dv 150)	left of $i_{21}$	right of $\underline{\text{nin}}$ deletion	>, 8%	$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 161)	left of $i_{434}$	right of $\underline{\text{nin}}$ deletion	>, 21%	$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 249)	between the left boundar- ies of $i_{21}$ and $i_{434}$	between $\underline{\text{P}}$ and $\underline{\text{nin}}$ deletion	>	$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 261)	left of $i_{21}$			$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 280)	left of $i_{434}$	right of $\underline{\text{nin}}$	>, 21%	$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 292)	in $i_{434}$	right of $\underline{\text{nin}}$ , left of $\underline{\text{Q}}$	>	$\lambda$ p gal 8 vir cII $\underline{\text{nin}}$
866 (dv 200)			21%	$\lambda$ gal 8 cI 857 $\underline{\text{nin}}$ at 42°C
866 (dv 204 LC) <sup>+</sup>			8% <sup>++</sup>	$\lambda$ gal 8 cI 857 $\underline{\text{nin}}$ at 42°C
866 (dv 204 DK) <sup>+</sup>				$\lambda$ gal 8 cI 857 $\underline{\text{nin}}$ at 42°C
866 (dv 309)	left of $i_{21}$			$\lambda$ gal bio 69 vir $\underline{\text{nin}}$



## Legend to Table 3

$\lambda$ dv Carrying Strains and Genetic Information about  
 $\lambda$ dv's Available

\*The left and right ends of  $\lambda$ dv, as determined by genetic recombination, are oriented according to the  $\lambda$  map (Davidson and Szybalski, 1971). The left boundary for  $i^{434}$  is 73.6, whereas that for  $i^{21}$  is given as 70.5 and 72.0 by Simon, Davis and Davidson (1971) and by Davidson and Szybalski (1971). In my hands, it is mapped at 71.0 (Fig. 2).

\*\* Physical size of  $\lambda$ dv monomers was tested by the method of compound phage formation as described in Experimental Procedure as well as in Results and Discussion.  $\approx$ , comparable to genetic size;  $>$ , larger than genetic size.

† Davidson and Szybalski (1971).

‡ Contains genes of galactose operon of about 8% of  $\lambda$  in size (Michael Feiss).

Δ Detailed genotype was not obtained from Dr. D. E. Berg.

ΔΔ The physical mapping data are not in agreement with the genetic analyses obtained from Dr. Berg. cf. Further Discussion.

( Table 3, continued )

§ KM 424 (rec A<sup>-</sup>, sus<sup>-</sup>, his<sup>-</sup>) carries  $\lambda$ dv 1 which is identical to the  $\lambda$ dv isolated by Matsubara and Kaiser (1968). 866 (rec A<sup>-</sup>, sus<sup>-</sup>, his<sup>-</sup>,  $\Delta$ gal -  $\lambda$ att) is a derivative of KM 424 which was first cured of  $\lambda$ dv 1 and then selected for a gal -  $\lambda$ att deletion. Both strains are slow growing with a generation time of 90 minutes or longer at 37° C in the medium used in this work.

+  $\lambda$ dv 204 DNA prepared in this laboratory ( $\lambda$ dv 204 LC) is different from that prepared in D. Kaiser's ( $\lambda$ dv 204 DK) due to a mix-up either in stabs or in DNA preparations. Both were then tested genetically by Berg.

++ These compound phages are at the limit of the packing capacity of the phage head (about 106%  $\lambda^+$ ) and are unstable. The plaques are mottled due to segregation.

to complete the cell lysis. Five ml of TM (TMG without gelatin) were added to each plate. The suspensions (with or without top agar) were collected after at least 2 hrs standing with occasional rocking. The phage suspensions were stored as such for infecting bacteria in liquid medium or further purified by banding in CsCl gradients (methods B (2) and A (5)) if they were to be used as DNA sources in heteroduplex mapping.

(3) Preparation of  $\lambda$  phage in broth.

An over-night culture (at 37° C) of E. coli K12 W3110 in trypton broth (1% trypton, 0.5% NaCl, pH 7.4) was used to inoculate 2 liters of broth containing 15 g trypton, 5 g yeast extract, 5 g NaCl per liter of water; pH was adjusted to 7.4 before autoclaving. 20 ml of 10% glucose, 5 ml of 1 M MgSO<sub>4</sub> and 0.2 ml of 0.5 M CaCl<sub>2</sub> were added after being sterilized separately. The culture was infected with the new phage stock at a multiplicity of infection of 0.05 when cell density reached  $2 \times 10^8$ /ml. The culture was kept at 37° C with good aeration until lysis (about 2-3 hrs). 20 ml of chloroform were then added and aeration was continued for another 10 min. The lysate was kept in the cold room overnight.

(4) Concentration of phage particles by the polyethylene glycol (PEG)-dextran sulfate (DS) two-phase separation method.

The lysate was adjusted to contain 2.9% NaCl, 1.5% PEG and 0.02% DS. The mixture was shaken vigorously in a separatory funnel and let stand in the cold room overnight for phase separation.

The lower phase (DS) and interphase (cake of cell debris) were discarded. Essentially all  $\lambda$  phage particles remained in the upper phase (PEG) under this condition. The solution was adjusted to be 8% in PEG and 0.05% in SD. The mixture was again vigorously agitated and placed in the cold room to allow the phases to separate. The lower phase was discarded. The interphase cake which contained practically all the phage particles was collected and centrifuged in a clinical centrifuge. The remaining PEG was removed with a pipet from the top, and DS was drained from the bottom of the centrifuge tube. The "phage cake" was suspended in 0.01 M Tris, 0.01 M  $\text{MgSO}_4$ , pH 8. To remove residual DS, 0.15 volume of 3 M KCl and solid DS were added to the suspension to give a DS concentration greater than 1%. The DS precipitation was removed by spinning in a clinical centrifuge for 10 min after standing at least 2 hrs at 4° C.  $\lambda$  phage recovered from the supernatant was greater than 98%.

(5) Banding of phage particles in preparative  $\text{CsCl}$  gradients.

To the phage suspension was added  $\text{CsCl}$  to a density of 1.5 gm/ml. Banding was carried out at 31 k or 37 k rpm for 48 hours in an SW 39 or SW50.1 rotor using a Beckman L, L2, or L2-65B centrifuge at room temperature. The phage particles were collected by dripping from a hole punched in the side of the centrifuge tube just below the visible phage band, and stored as such at 4° C.

## B. Preparation of $\lambda$ phage on agar plates.

- (1) The phages were grown as in A (1) and (2).
- (2) Purification of phage particles.

The phage suspension ( $\pm$  top agar) was spun in a Sorvall SS 34 rotor at 10 k for 5 min to remove cell debris (and agar). The phage particles were pelleted by spinning at 17 k for 2.5 hr in an SS 34 rotor at 4° C. The pellet was resuspended in 0.05 M Tris, 0.02 M  $\text{MgSO}_4$ , pH 8. Low-speed centrifugation was repeated once more. To the phage suspension was then added CsCl and banded as described in A (5).

## C. Preparation of compound phages on agar plates.

- (1) Growth of 829 S.

Over-night cultures of 829 S were grown in TYM (10 g trypton, 2 g yeast extract, 5 g NaCl, 2 g  $\text{MgSO}_4 \cdot 6\text{H}_2\text{O}$  per liter of water, supplemented after autoclaving with 10 ml of sterilized 20% maltose) and stored as such at 4° C for future use.

- (2) Preparation of phage.

Fresh TYM was inoculated with the over-night culture and was aerated at 42° C till  $A_{600} = 0.8$ . Duplicate phages received from Dr. D. E. Berg were used without single-plaque purification. The phage stocks were diluted with TMG and adsorbed to the young 829 S culture at 37° C for 10 min with shaking. The infected bacteria were plated on agar plates to give about 5000 plaques/plate. Top agar (5 ml/plate) contained 1% trypton, 0.75% agar, 0.5% NaCl, and 0.2%  $\text{MgSO}_6 \cdot \text{H}_2\text{O}$ . Bottom agar was the same as in A (1). The plates were kept in a 37° C oven for 5 hr, then were overlaid with 5 ml/plate

of 0.05 M Tris, 0.02 M  $\text{MgSO}_4$ , pH 8. The suspension and the top agar were collected, and 5 ml of chloroform were added per 20 plates.

### (3) Purification of compound phage particles.

The phage particles were first concentrated by low-and high-speed centrifugation as in B (2). Then 5 ml of phage suspension were laid on top of a step gradient consisting of 1 ml of  $\rho = 1.6 \text{ gm/ml}$   $\text{CsCl}$ , 1.5 ml of  $\rho = 1.5 \text{ gm/ml}$   $\text{CsCl}$ , 1 ml of  $\rho = 1.4 \text{ gm/ml}$   $\text{CsCl}$  and 4 ml of 10% sucrose. (All were buffered with 0.05 M Tris, 0.02 M  $\text{MgSO}_4$ , pH 8.) The gradient was spun at 37 k for 4 hr at  $20^\circ \text{C}$  in an SW41 rotor using a Beckman L2-65B centrifuge. Two bands were visible containing duplicate phages, segregants and a few triplicate phages in the case of compound phage containing  $\lambda\text{dv}1$ . There were three not well-resolved bands containing duplicate phage, segregants,  $i^\lambda$  recombinants, and duplicate phages with deletion in preparations of compound phages containing  $\lambda\text{dv}249$  and  $\lambda\text{dv}292$ . A hole was punched below the lowest band and the phage particles were collected separately whenever possible.

## 5. Preparation of closed circular $\lambda\text{dv}$ DNAs.

### (1) Growth of carrier cells.

A culture of 250 ml, 500 ml or 1 liter was grown from an overnight culture which was inoculated either directly from a stab or from a tested, purified single colony (see legend for Table 4 for details) at  $37^\circ \text{C}$  with good aeration until  $A_{590}$  reached 1.0 in a 50-fold diluted Vogel-Bonner medium. The medium was supplemented with 0.4% glucose, 0.4% Casamino acid, 250  $\mu\text{g/ml}$  adenosine, and

0.0005% of thiamine·HCl which were sterilized separately. Vogel-Bonner medium contained 10 gm  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ , 100 gm citric acid· $\text{H}_2\text{O}$ , 655 gm  $\text{K}_2\text{HPO}_4 \cdot 3 \text{H}_2\text{O}$ , 175 gm  $\text{NaNH}_4\text{HPO}_4 \cdot 4 \text{H}_2\text{O}$ , and 515 ml distilled water to give 1 liter of pH 7 solution.

(2) Isolation of closed circular  $\lambda$ dv DNA.

The method developed by Sharp, Hsu, Ohtsubo and Davidson (1972) for episome isolation was adopted.

The cells were pelleted in a Sorvall GSA rotor at 5 k rpm for 15 min at 4° C, then washed twice with cold TES buffer (0.05 M Tris, 0.005 M EDTA, 0.05 M NaCl, pH 8).

The cells from 1 liter of culture were resuspended in 50 ml of cold TES with 1 mg/ml lysozyme, 0.1 mg/ml RNase, and 10% sucrose and incubated at 37° C for 10 min, then chilled over ice. The spheroplasts were lysed by adding a half volume of 2% sarcosyl (in water). The host chromosomal DNA was sheared in 2 batches through a 50 ml plastic disposable syringe for 6 times at a speed of about 5-10 sec/solume. Linear DNA was denatured by adding 1 N NaOH to pH 12.3 with vigorous stirring. The lysate was maintained at pH 12.3 for 3 min, then neutralized with concentrated Tris-HCl to pH 8.5 after adding enough 5 N NaCl to give  $[\text{Na}^+] = 0.3 \text{ N}$ . Most of the denatured DNA was removed by adding 30 g of nitrocellulose which had been ground and prewashed twice with TES buffer. The mixture was rotated gently at 4° C for 45 min. Nitrocellulose was then removed by low-speed centrifugation and filtering through glass wool. To pellet the  $\lambda$ dv DNA, the supernatant was spun in an SW 25.2

rotor with a 2 ml cushion of  $\rho = 1.7$  gm/ml CsCl in TES at 4° C, using a Beckman L2 or L2-65B centrifuge.  $\lambda$ dv DNA was recovered in the bottom 8 ml of the solution. To the DNA solution was then added solid CsCl according to the formula

$$\text{gm CsCl} = V (1.55 - \rho_0)/0.61$$

where  $\rho_0$  was the density of the solution before adding CsCl, and V the volume of the solution before adding CsCl. Enough stock solution (10 mg/ml in water) of ethidium bromide was added to give a concentration of 400  $\mu$ g/ml. DNA was then banded at 37 k rpm for 48 hr at 20° C in an SW 50.1 rotor. Usually 2 or 3 bands were visible under a UV light. The bottom, middle and top bands contained closed circular  $\lambda$ dv DNA, denatured DNA and linear native DNA, respectively. The bands were collected separately from the bottom of the centrifuge tube. The top (plus the middle) and bottom bands were pooled separately (2 to 2.5 ml each) and rebanded for 16 hr at 20° C, 37 k rpm in an SW 50.1 rotor. The bottom band was then collected and pooled. Ethidium bromide was removed by dialyzing either against large volumes of 0.01 M Tris, 0.001 M EDTA, 0.5 M NaCl pH 8 at 4° C or against small volumes of 0.025 M Tris, 0.005 M EDTA, 0.8 M NaCl pH 8.5 containing 0.05 g/ml Dowex-50. Dowex-50 was pretreated as follows: 25 g of Dowex-50 was suspended in water, titrated to pH 12.3 with NaOH, neutralized to pH 8.5 with HCl, then was washed several times with buffer containing 0.05 M Tris 0.01 M EDTA, 1.3 M NaCl, pH 8.5, and finally suspended in 250 ml of the same buffer. An equal volume of 0.3 N NaCl was added when used for dialysis. Occasionally, traces of ethidium bromide remained bound to DNA



even after prolonged dialysis against Dowex-50. However, it did not interfere with electron microscopy. The DNA solution was then dialyzed against 0.01 M Tris, 0.001 M EDTA, pH 8 and recovered and stored as such at 4° C. The DNA was examined in an electron microscope for sizing and for purity checking. When contaminating double-or single-stranded linear DNA was more than 5%, ethidium bromide-CsCl banding was repeated once more.

#### 6. Nicking of closed circular $\lambda$ dv DNA.

$\lambda$ dv DNA solution in 0.05 M Tris, 0.01 M EDTA, 0.1 N NaCl and 50  $\mu$ g/ml ethidium bromide was put in a stoppered glass curvett and was stirred gently over a magnetic stirrer. Light from a Kodak slide projector with a 500 w bulb was projected onto the curvett at a distance about 20 cm with a piece of aluminum foil behind as a reflector. A fan was used to dissipate the heat. Illumination of 10 min and 25 min introduced about 0.5 single-stranded nicks per molecule for molecules of sizes about 30% and 7% of wild type  $\lambda$  DNA, respectively. The DNA solution was then dialyzed into 0.01 M Tris, 0.001 M EDTA, pH 8 for storage at 4° C.

#### 7. Self renaturation of $\lambda$ dv and hybridization of $\lambda$ dv and $\lambda$ DNAs.

$\lambda$  phage particles in CsCl and nicked  $\lambda$ dv DNA were used as DNA sources. The procedures for self renaturation and for making heteroduplexes are described in Davis, Simon and Davidson (1971) and in Chow, Boice and Davidson (1972). The DNA concentrations used were 2.5  $\mu$ g/ml of  $\lambda$ dv DNA and 1-2.5  $\mu$ g/ml of  $\lambda$  DNA. Renaturation in 50% formamide by either adding formamide to the DNA solution

or by dialyzing against buffered (with 0.1 M Tris, 0.01 M EDTA, pH 8.6) 50% formamide was carried out at room temperature for 1 to 2 hr. Renaturation was stopped by dialyzing against 0.01 M Tris, 0.001 M EDTA, pH 8 at 4° C.

#### 8. Snap-back experiment.

The nicked  $\lambda$ dv DNA was denatured and reneutralized. It was then mounted immediately using the formamide or aqueous technique (Davis *et al.*, 1971; Chow *et al.*, 1972; Chow and Davidson, 1973), or mounted after dialyzing into 0.01 M Tris, 0.001 M EDTA, pH 8 at 4° C.

#### 9. Electron microscopy.

(1) The formamide technique was used for sizing and mapping. The mounting conditions were: from a spreading solution of 40% or 45% formamide onto a hypophase of 10% or 17% formamide, respectively.

(2) For partial or full denaturation of DNA containing snap-back regions, high formamide mounting was used:

(a) 85% formamide onto 50% formamide at 40° C.

Nicked dv 104 was denatured in 0.125 N NaOH for 10 min at room temperature. The spreading solution (per 100  $\mu$ l) was made up by adding to 2.5  $\mu$ l of the alkaline  $\lambda$ dv solution the following solutions: 1.25  $\mu$ l of 2 M Tris (1.8 M Tris-HCl, 0.2 M Tris-OH) pH 7.1, 3.5  $\mu$ l of 1.4 M Tris, 0.06 M EDTA, pH 8.6, 5  $\mu$ l of cytochrome c at 1 mg/ml in 1 M ammonium acetate, 2.75  $\mu$ l of  $\phi$ X174 at 5  $\mu$ g/ml and 85  $\mu$ l of formamide. The hypophase contained 0.7 mM Tris, 0.03

m M EDTA pH 8.6 and 50% formamide.

(b) 90% formamide onto 55% formamide at 40° C.

Spreading solution (per 100  $\mu$ l) contained: 2.5  $\mu$ l of the alkaline  $\lambda$ dv solution, 1.25  $\mu$ l of 2 M Tris, pH 7.1, 2.25  $\mu$ l of 1.4 M Tris, 0.06 M EDTA, pH 8.6, 1  $\mu$ l of cytochrome c at 5 mg/ml in 5 M ammonium acetate, 3  $\mu$ l of  $\phi$ X174 at 5  $\mu$ g/ml, and 90  $\mu$ l formamide. Hypophase was the same as in (a) except that the formamide concentration was increased to 55%.

(3) For obtaining partial denaturation maps of  $\lambda$  and  $\lambda$ dv DNAs, the spreading solution contained 4.2 m M Tris, 1.6 m M EDTA, 40 m M ammonium acetate, 40  $\mu$ g/ml cytochrome c, pH 8.6 and 85% formamide at room temperature. Hypophase contained 0.37 m M Tris, 0.16 m M EDTA, pH 8.6 and 50% formamide at room temperature.

### Methods for Mapping $\lambda$ dv DNAs

A  $\lambda$ dv DNA was mapped by analyzing heteroduplexes of  $\lambda$ dv with wild type  $\lambda$  and/or with mutant  $\lambda$  DNAs. The physical size of the  $\lambda$ dv DNA was obtained by length measurements either on the heteroduplexes or on the open circular molecules. The methods used are explained in detail below.

- A. Hybridization with the "parental" phage DNA to give a heteroduplex without internal reference points. (For example,  $\lambda$ dv's 1 and 21 with wild type  $\lambda$ ;  $\lambda$ dv's 154 and 161 with  $\lambda$  p gal 8 vir cII nin, etc.; see Table 3 for parental phage.) (Fig. 1 (a) and (b)).

$(100 - r)$  and  $(100 - r - D_1)$  (in % of  $\lambda$ ) give the right and left end points of the  $\lambda$ dv;  $D_1$  is the physical length of the unique genetic sequence, and  $(S + D_1)$  is the physical length of the complete  $\lambda$  dv. The accuracy of this method depends upon the intactness of the  $\lambda$  DNA, the homogeneity of the  $\lambda$ dv, and the accuracy of measurement of the single stranded segment,  $r$ , which is rather long in some cases.

- B. Hybridization with  $\lambda$  DNAs to give heteroduplexes with internal reference points.

B-1. Hybridization of  $\lambda$ dv's derived from  $\lambda$  nin<sup>+</sup> phage with both (a) wild type ( $\lambda$  c<sub>26</sub>) (Fig. 1 (a) and (b)) and (b)  $\lambda$  b<sub>2</sub>b<sub>5</sub>c DNAs (Fig. 1 (c), (d), and (h)).

$\lambda$  b<sub>5</sub> is identical to  $\lambda$  i<sub>21</sub>; thus the left end point of  $D_2$  is the right end point of i<sub>21</sub> and is mapped in my hands to be at 79.6 from

Fig. 1 Heteroduplex structures between  $\lambda$ dv and  $\lambda$  DNA's.

- (a) Wild type  $\lambda$  DNA (  $\lambda c_{26}$  ) / circular  $\lambda$ dv from  $\lambda nin^+$  phage.
- (b)  $\lambda c_{26}$  / linear  $\lambda$ dv from  $\lambda nin^+$  phage.
- (c)  $\lambda b_2 b_5 c$  / circular  $\lambda$ dv from  $\lambda nin^-$  phage when  $\lambda$   $\lambda$ dv does not extend the left boundary of  $i^{21}$ .
- (d) and (h)  $\lambda b_2 b_5 c$  / circular  $\lambda$ dv from  $\lambda nin^+$  phage when  $\lambda$   $\lambda$ dv extends to the left of  $i^{21}$ .
- (e)  $\lambda c_{26}$  / circular  $\lambda$ dv from  $\lambda nin^-$  phage.
- (f)  $\lambda c_{26}$  / circular  $\lambda$ dv trimer from  $\lambda nin^-$  phage.
- (g)  $\lambda c_{26}$  / linear  $\lambda$ dv oligomer from  $\lambda nin^-$  phage.

(f) and (g) are different versions of (e), whereas (h) that of (d). The genes (  $cIII$ ,  $i^{21}$ ,  $O$ , and  $P$  ) on the  $\lambda$  DNA are oriented according to the  $\lambda$  map ( Davidson and Szybalski, 1971), ——— ,  $\lambda$ DNA; ——— ,  $\lambda$ dv DNA;  $D_1$ ,  $D_2$ ,  $D'_2$ ,  $D_3$ , and  $D_4$  are the duplex regions; S is the single stranded portion of the  $\lambda$   $\lambda$ dv ;  $r$  stands for the single-stranded  $\lambda$  DNA starting from the right end of the duplex  $D_1$ ,  $D_2$ , or  $D_4$  to the right terminus of the  $\lambda$  DNA; ( $D_3$ ) and ( $D_4$ ) designate the single strands of  $\lambda$ dv corresponding in sequence, and in length to  $D_3$  and  $D_4$  respectively;  $nin^+$  represents the sequence deleted in the  $nin$  deletion.

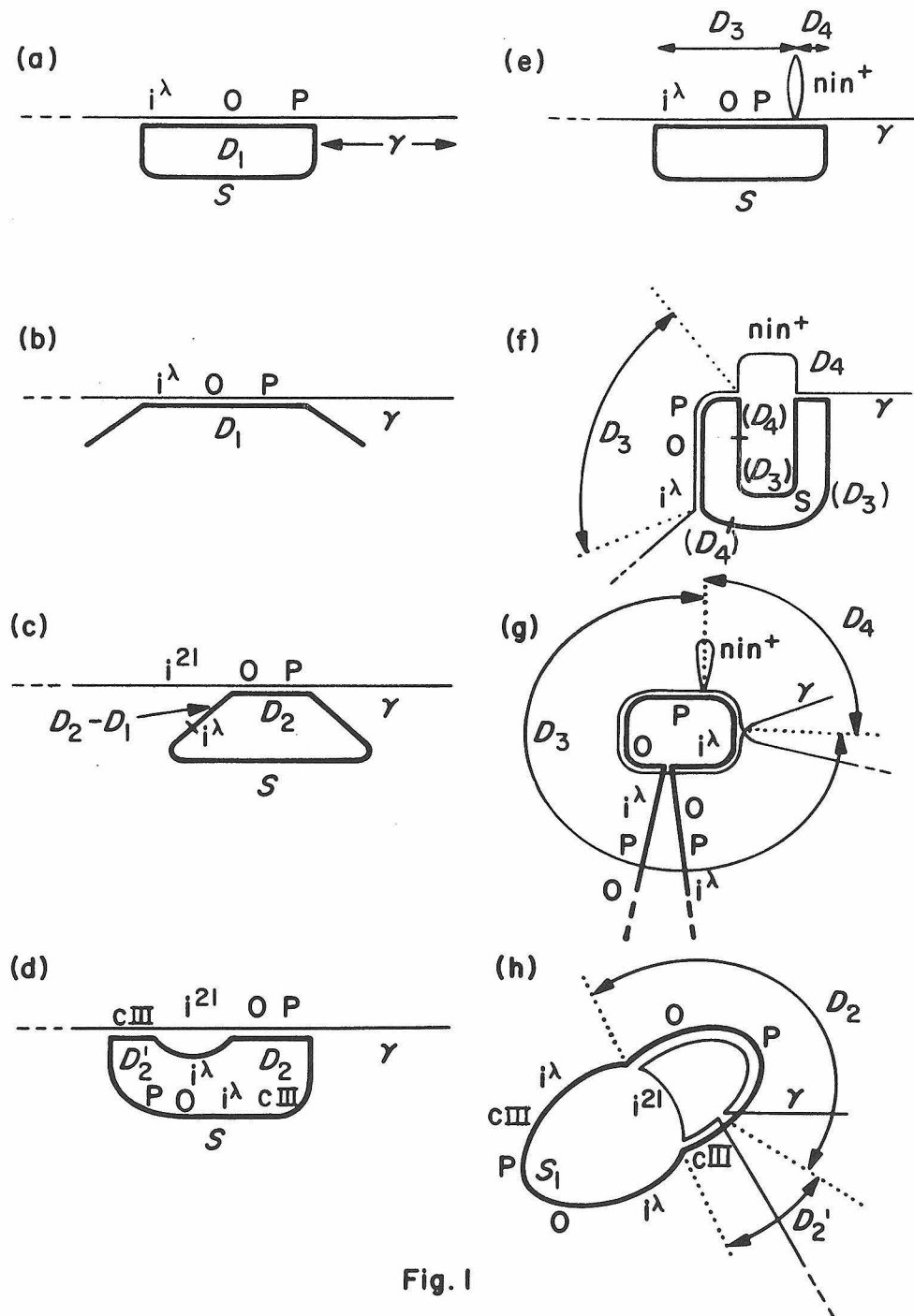


Fig. 1

heteroduplex between  $\lambda_{b_2b_5c}$  and  $\lambda_{p\ gal\ 8\ vir\ cII\ nin}$  (Fig. 2) (Plate I) as opposed to 79.8 by Davidson *et al.* (1971) and 79.5 by Simmon *et al.* (1971).  $79.6 + D_2$  = the right end point;  $79.6 + D_2 - D_1$  = the left end point when the  $\lambda_{dv}$  DNA does not extend beyond the left boundary of  $i^{21}$  (Fig. 1 (c)), or  $71 - D_2'$  = the left end point when  $\lambda_{dv}$  DNA does extend beyond  $i^{21}$  (Fig. 1 (d) and (h)). The left boundary of  $i^{21}$  maps at 71.0 in my hands (Fig. 2). However, it is given as 72 and 70.5, respectively, by the other two groups of investigators above. Due to the differences in the mapping positions for the boundaries of  $i^{21}$ , the important figures for  $\lambda_{dv}$  mapping are then  $D_1$ ,  $D_2$ , and  $D_2'$  instead of the value calculated as above. Examples are presented in Plates II and III.

B-2. Hybridization of  $\lambda_{dv}$ 's derived from  $\lambda_{nin^-}$  phages with wild type  $\lambda$  DNA ( $\lambda_{c_{26}}$ ).

Heteroduplex of  $\lambda_{p\ gal\ 8\ vir\ cII\ nin}$  with  $\lambda_{c_{26}}$  shows  $nin$  as a deletion extending from 83.8 to 89.6% of  $\lambda$  (as compared to 83.8 to 89.2% according to Davidson *et al.*, 1971). The heteroduplex structure of most of the  $nin^-$   $\lambda_{dv}$ 's with  $\lambda_{c_{26}}$  are shown in Fig. 1 (e).  $(89.6 + D_4)$  and  $(83.8 - D_3)$  give the right and the left end points of the  $\lambda_{dv}$ . Obviously, the  $nin^+$   $\lambda_{dv}$  can also be mapped by this method when heteroduplexes are made between these  $\lambda_{dv}$ 's and  $\lambda_{nin^-}$  DNA. For  $nin^-$   $\lambda_{dv}$  that does not extend beyond the  $nin$  deletion, mapping method A was used.

Fig. 2 Mapping of  $\lambda$  p gal 8 vir nin -- the parental phage of  $\lambda$  dv's containing inverted repeats.

————,  $\lambda$  DNA; -----, deletion; - · - · -, substitution deletion. The  $\lambda^{att}$  site maps at 57.1 in both heteroduplexes. It is normalized to 57.3 ( cf. the  $\lambda$  map ). The other segments are then normalized accordingly.



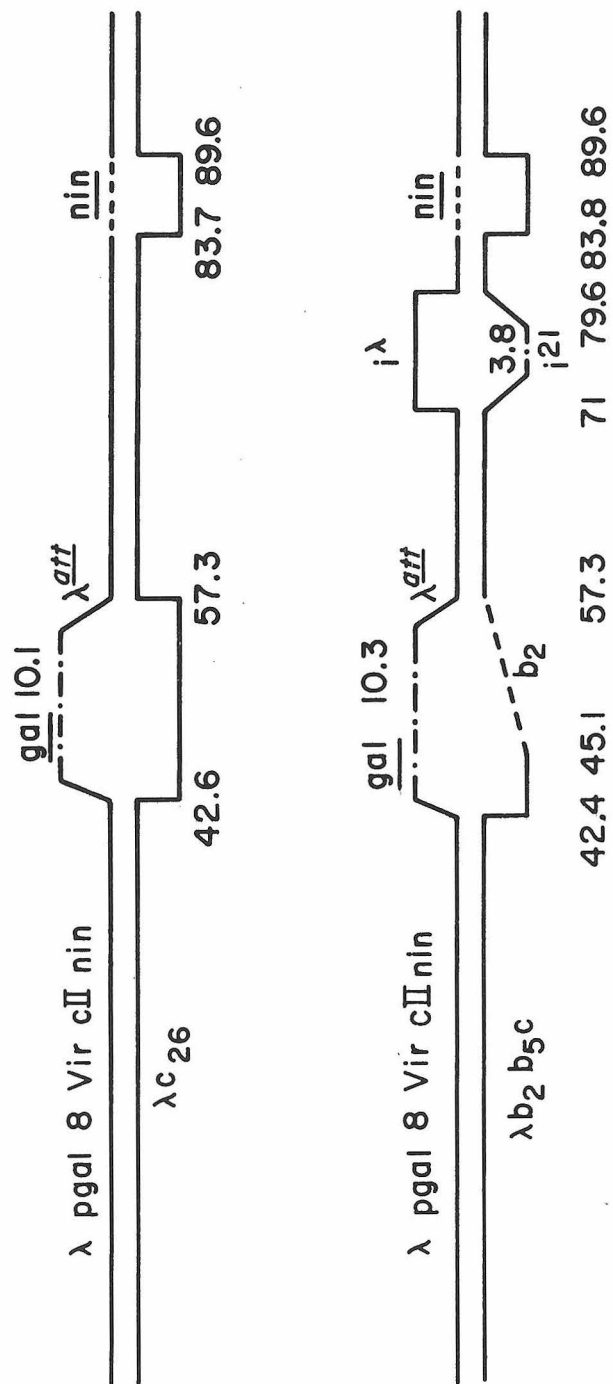


Fig. 2

Plate I Heteroduplex of  $\lambda$ b<sub>2</sub>b<sub>5</sub>c and  $\lambda$ p gal 8 vir cII nin mounted from 45 % onto 17 % formamide.

The single strands are labeled. The small circular molecules are  $\phi$  x 174 RF II added as measurement standard. A schematic representation is seen in Fig. 2.

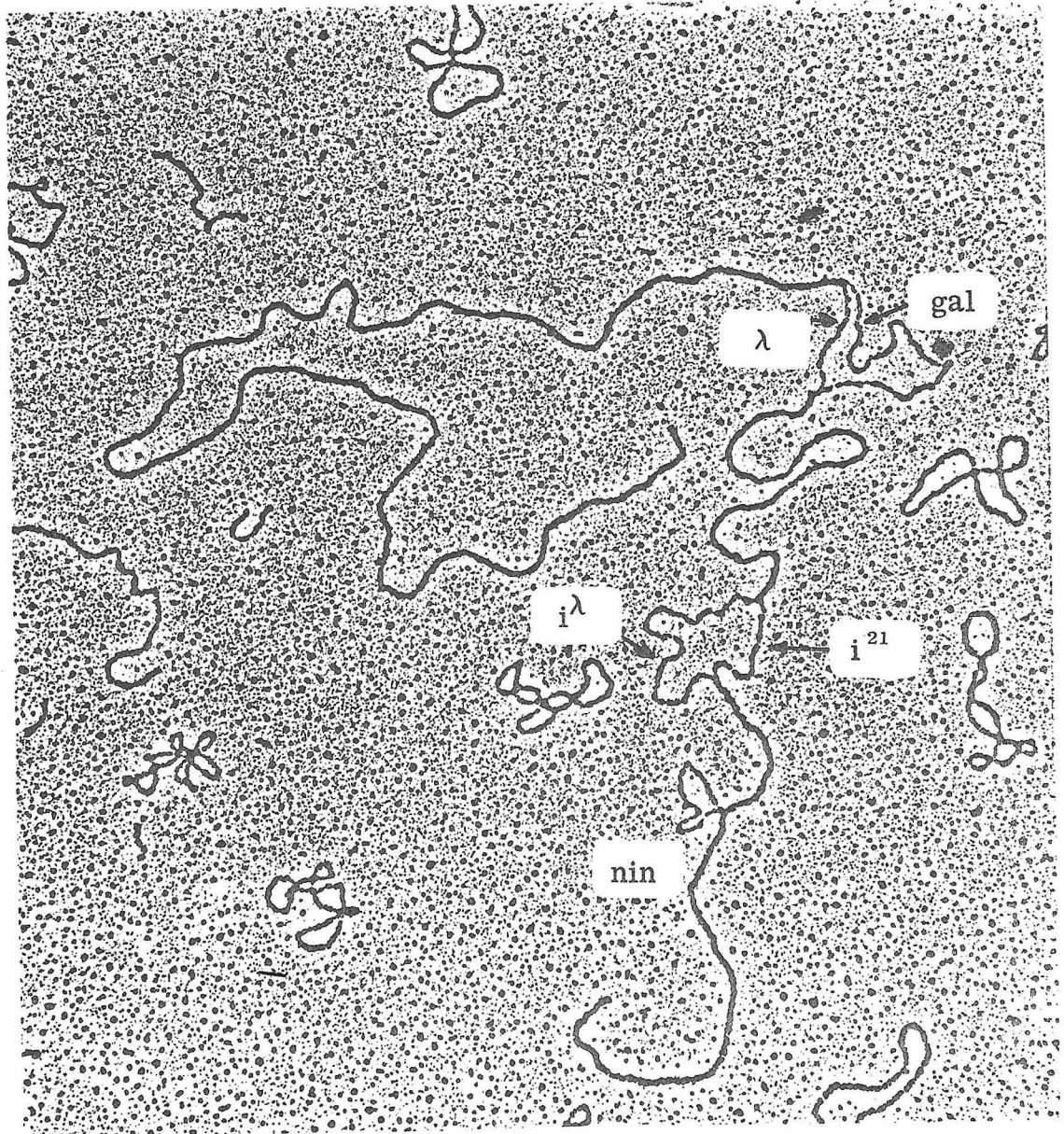


Plate I

Plate II Heteroduplexes of  $\lambda$  and  $\lambda$ dv 1 DNA's.

A  $\lambda\lambda_{c_{26}} / \lambda$ dv 1 circular dimer.

B  $\lambda b_2 b_5 c / \lambda$ dv 1 circular dimer.

Both heteroduplexes were mounted from 40 % onto 10 % formamide. Schematic representations are seen in Fig. 1(a) and (c).  $\blacktriangle$  points to the junctions of  $\lambda$  and  $\lambda$ dv DNA's, i. e., the junction of duplex and single strands. The small circles are  $\phi$ X 174 and  $\phi$ X 174 RF II.

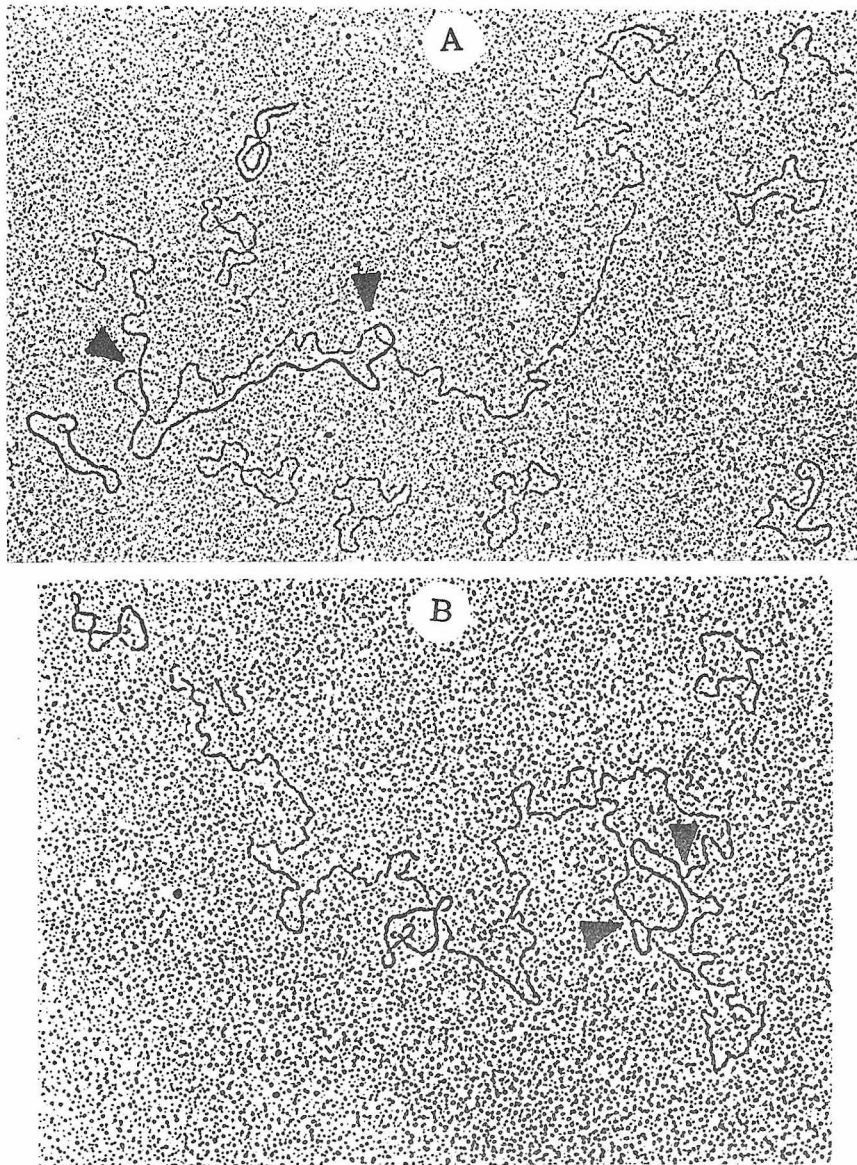


Plate II

Plate III Heteroduplexes of  $\lambda$  and  $\lambda$ dv 203 DNA's.

A  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$  /  $\lambda$ dv203 circular dimer. This molecule is depicted in Fig. 1(d).

B  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$  /  $\lambda$ dv 203 linear oligomer. This molecule is depicted in Fig. 1 (h) with S broken.

C  $\lambda \underline{c}_{26}$  /  $\lambda$ dv 203 circular dimer. A schematic representation is seen in Fig. 1 (a).

Some of the duplex and single strands are labeled.  $\blacktriangle$  points to the junction of  $\lambda$  and  $\lambda$ dv DNA's . The molecules were mounted from 45% onto 17 % formamide.

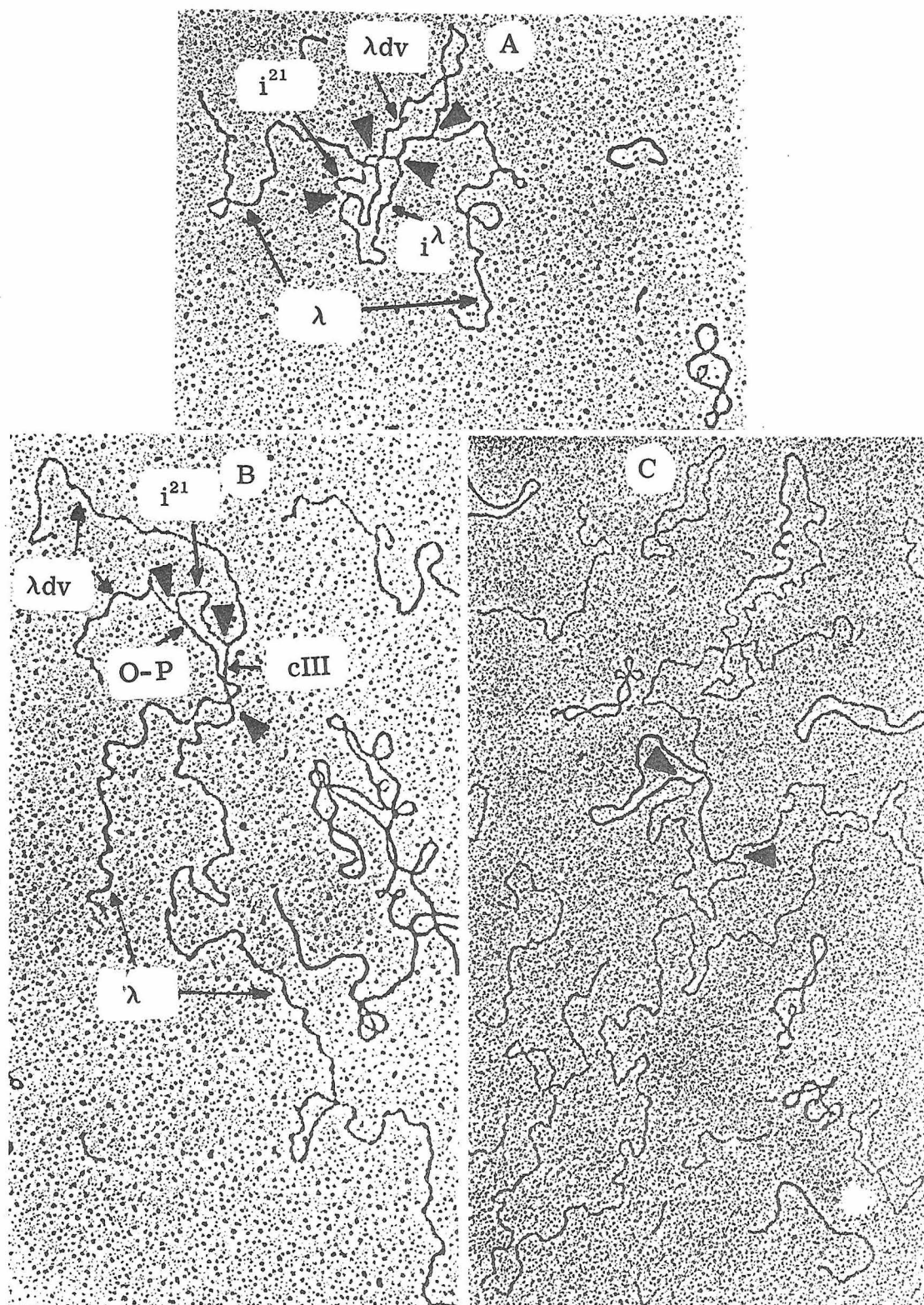


Plate III



Since  $\lambda$ dv's can exist as dimers and higher oligomers, the heteroduplex structures with  $\lambda$  DNAs are not always as simple as those depicted in Fig. 1 (a) to (e). More complicated structures are often encountered. A few examples are presented in Fig. 1 (f) to (h) and in Plates III B, IV A and B. (Other more complicated structures of these types have also been observed.) Structures 1 (f) and (g) are different versions of category 1 (e) heteroduplex. The  $\lambda$ dv strand of 1 (g) can branch migrate everywhere around the duplex circle except at the point where the  $\lambda$  strands branch out (It then becomes identical to 1 (e) ). Structure 1 (h) is of category 1 (d). As indicated in the figure, data analyses for the more complicated heteroduplexes are the same as those for the simple ones.

### C. Mapping of $\lambda$ dv via heteroduplexes with compound phages.

$\lambda$ dv carrier cells are not immune to  $\lambda$ i<sup>21</sup> (Berg, 1971). When they are infected with phage  $\lambda$ i<sup>21</sup>,  $\lambda$ dv DNA can be inserted into the  $\lambda$ i<sup>21</sup> DNA by recombination to give a DNA molecule containing, in tandem, two or more copies of the genes contained in the  $\lambda$ dv (Kellenberger - Gujer, 1971; Berg, 1971; Berg *et al.*, 1972). The generation of a duplicate phage by recombination between a tandem  $\lambda$ dv dimer and the superinfecting phage  $\lambda$ i<sup>21</sup> DNA is shown in Fig. 3 (a) and (b). The formation of a triplicate phage is shown in Fig. 3 (c). Presumably, the latter can also be formed by inserting a  $\lambda$ dv monomer into a duplicate phage DNA. Quintuplicate phage with four copies of  $\lambda$ dv21 (6.8% of  $\lambda^+$ ) inserted into  $\lambda$ b<sub>538</sub>i<sup>21</sup> sus  $\underline{P}$  has been detected on CsCl density gradients (Berg *et al.*, 1972). Duplicate



Fig. 3 Compound phage formation between  $\lambda$ dv and superinfecting

$\lambda$ i<sup>21</sup> phage DNA's ( Berg, et al., 1972 )

(a) and (b) generation of duplicate phages

(c) generation of triplicate phage.

Notation is the same as in Fig. 1

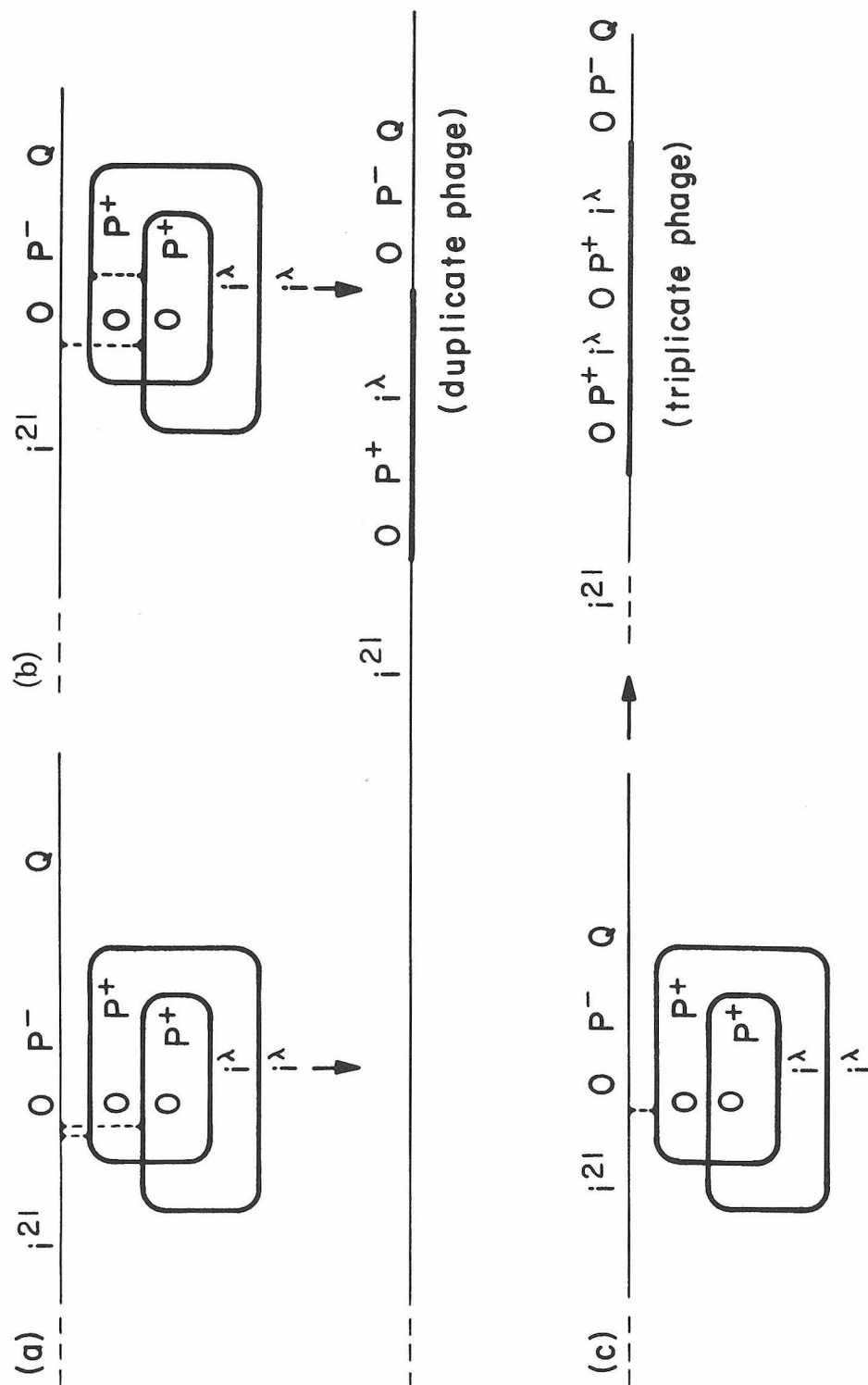




Fig. 3

Plate IV Heteroduplexes of  $\lambda_{c_{26}}$  and  $\lambda_{dv}$  204DK DNA's.

The schematic representations of A and B are seen in Fig. 1 (f) and (g) respectively ; that of C in Fig. 1 (e) with **S** broken. For notation, see Fig. 1.  points to the junction of  $\lambda$  and  $\lambda_{dv}$  DNA's.  points to the nin deletion. The molecules were mounted from 45 % onto 17 % formamide.

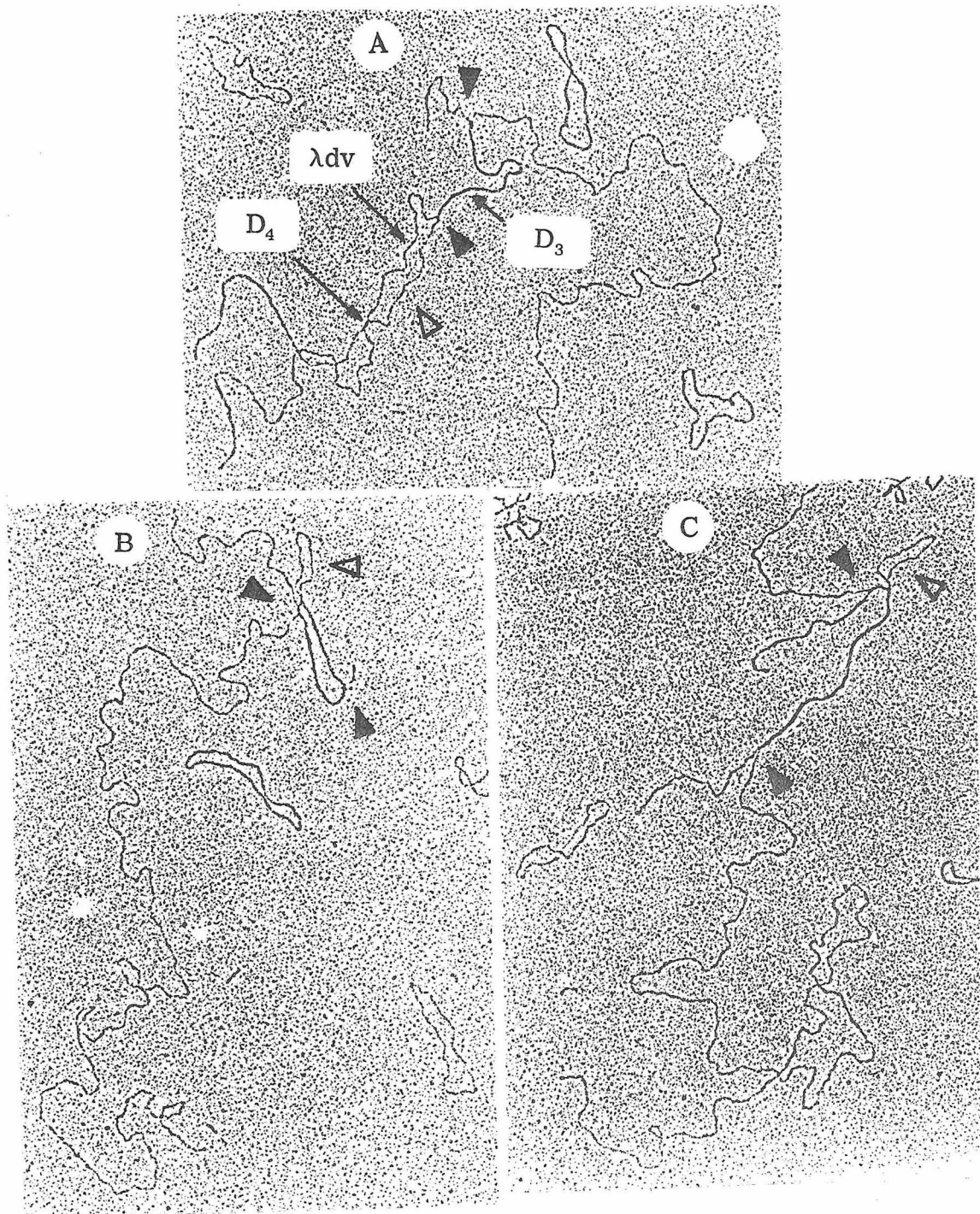


Plate IV

and triplicate phages containing one and two copies of  $\lambda$ dv1 in  $\lambda$ b<sub>538</sub> i<sup>21</sup> sus P have been observed by electron microscope. Examples are presented in Plate V A, C and D. Since only about 106% of wild type  $\lambda$  DNA can be packed into the  $\lambda$  phage head (Berg *et al.*, 1972), deletion mutants of  $\lambda$  i<sup>21</sup> are used as the super-infecting phage to accommodate the incoming  $\lambda$ dv.  $\lambda$ b<sub>538</sub> i<sup>21</sup>, with a 21% deletion, was used in my electron microscope mapping work. When the duplicate phage is hybridized to various  $\lambda$  DNA's, heteroduplexes such as in Fig. 4 (a), (b), and (c) can be seen. The single-stranded loop i <sup>$\lambda$</sup>  - O - P in Fig. 4 (a) and (b) represents the physical size of the inserted  $\lambda$ dv. The loop migrates between  $\alpha$  and  $\beta$ , the end points of the duplicated segment. Loop migration due to tandem duplication of genes has been observed in derivatives of  $\lambda$  phage (Bellett, Busse and Baldwin, 1971; Busse and Baldwin, 1972). The distance between  $\alpha$  and  $\beta$  can also be measured as the length of the single-stranded loop i<sup>21</sup> - O - P in Fig. 4 (c) minus the length of i<sup>21</sup> (Fig. 2). The segment between  $\alpha$  and  $\underline{r}$  represents the genetic content of the  $\lambda$ dv DNA. Examples of these heteroduplexes are shown in Plate V A, B and C.

Fig. 4 Heteroduplexes between  $\lambda$  and compound  $\lambda$  phage DNA's.

(a), (b) and (c) are drawn according to heteroduplexes observed in hybridization preparations of  $\lambda b_{538} i^{21} - \lambda dv 1$  and different  $\lambda$  DNA's. The single-stranded segment  $\gamma\delta$  is present because  $\lambda dv 1$  does not contain the complete  $i^\lambda$  region.

(d) and (e) depict heteroduplexes of DNA's of  $\lambda$  and compound phage  $\lambda$  containing  $\lambda dv$  with inverted repeat. Further discussion is presented in the text.

Notation is the same as in Fig. 1.  $\underline{P}'$  and  $\underline{Q}'$  are complementary sequences of  $\underline{P}$  and  $\underline{Q}$ .

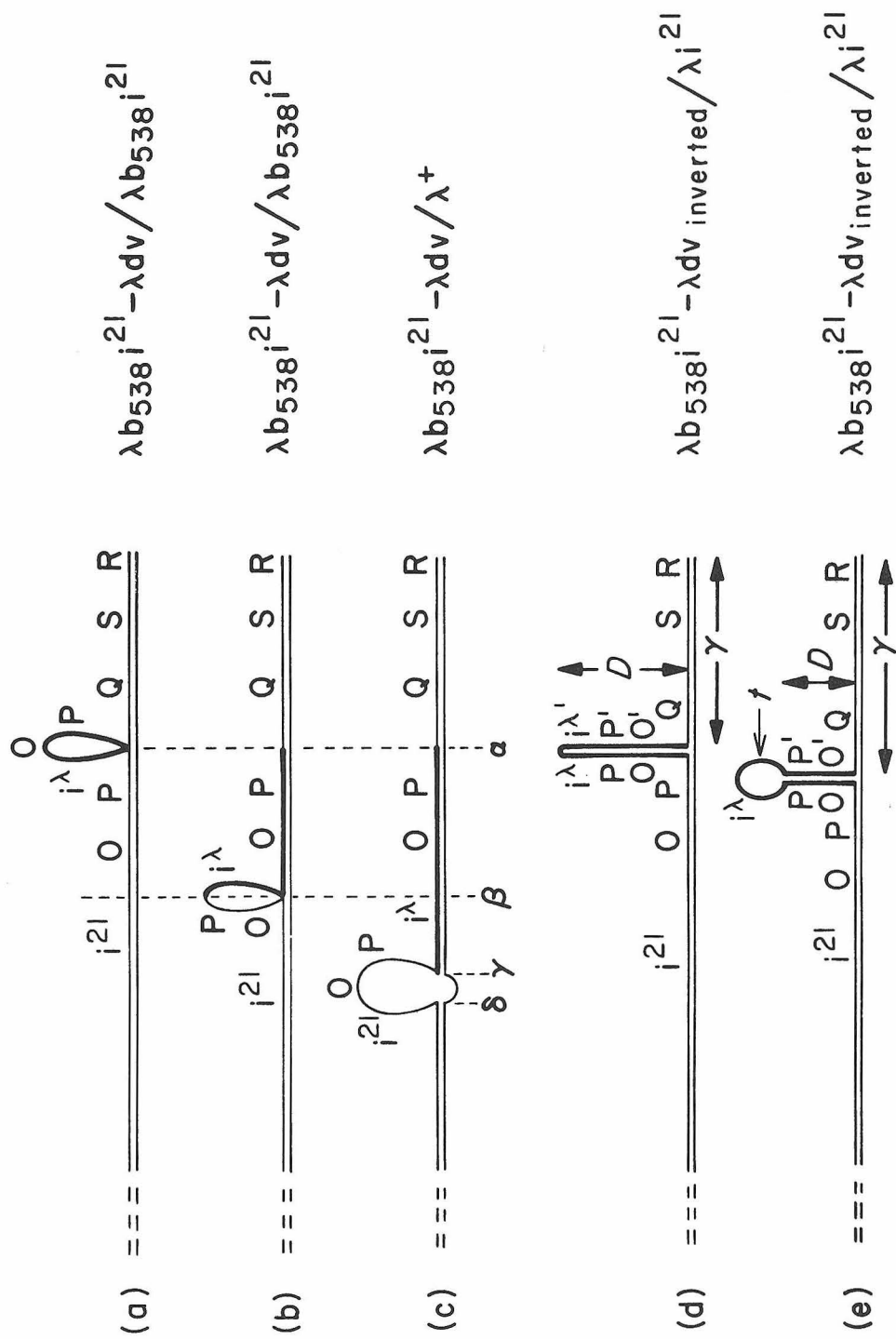


Fig. 4

Plate V Heteroduplexes of DNA's of  $\lambda$  and compound phages containing  $\lambda$ dv 1.

- A and C  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 1 /  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$ .  $\blacktriangle$  points to the inserted  $\lambda$ dv 1 monomer;  $\triangle$  points to the small deletion loop formed due to the difference between deletions  $\underline{b}_{538}$  and  $\underline{b}_2$ . A partial schematic representation is seen in Fig. 4 (a) and (b). Molecule C also shows the pulled-apart  $\lambda$  cohesive ends.
- B  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 1 /  $\lambda \underline{c}_{26}$ .  $\blacktriangle$  points to the inserted  $\lambda$ dv 1 monomer;  $\triangle$  points to the  $\underline{b}_{538}$  deletion loop. For a partial schematic representation, see Fig. 4 (c).
- D  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 1 -  $\lambda$ dv 1 /  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$ .  $\blacktriangle$  points to the inserted  $\lambda$ dv 1 dimer;  $\triangle$ , the same as in A and C. The molecules were mounted from 45 % onto 17 % formamide.



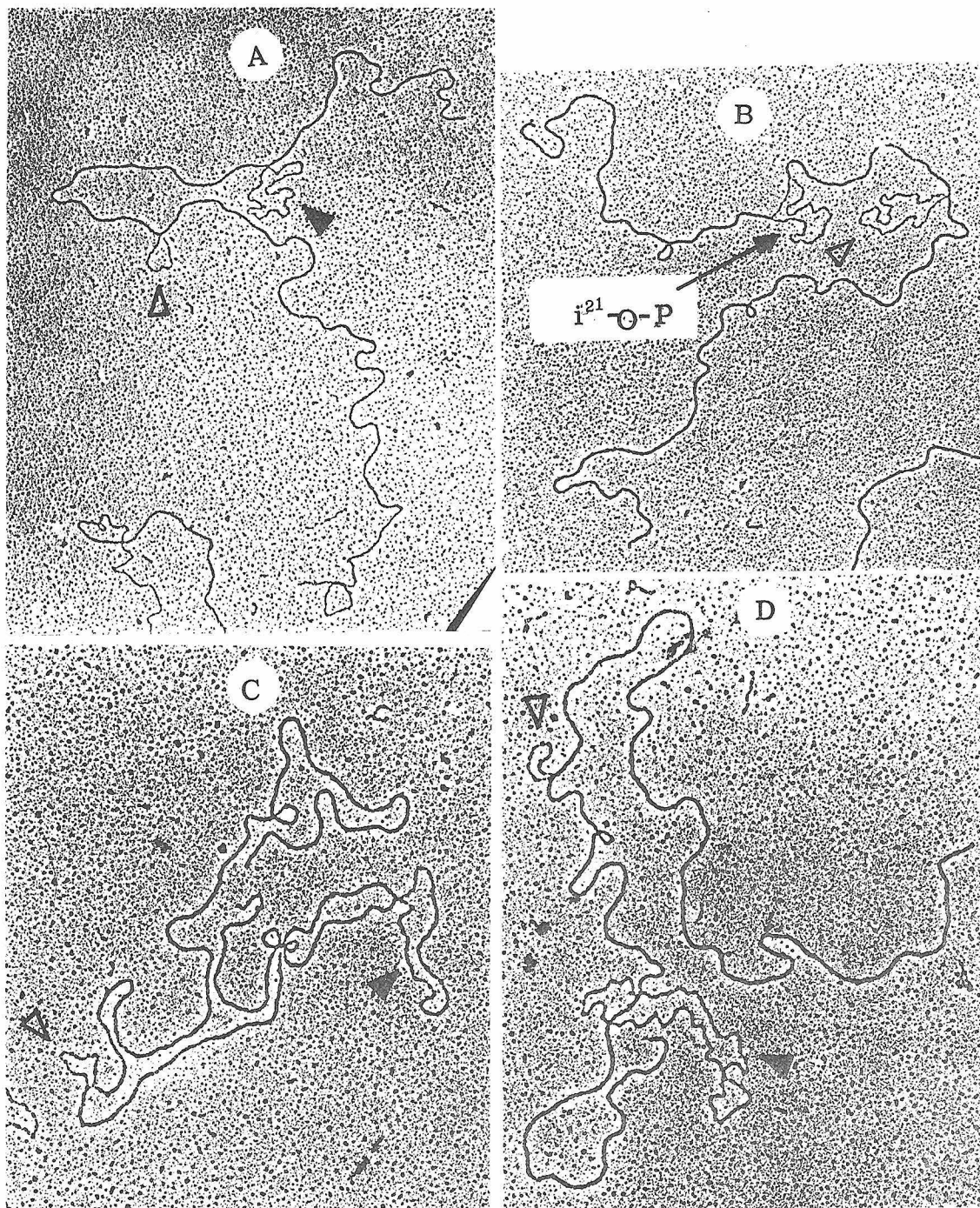


Plate V

## Results and Discussion

### 1. $\lambda$ dv's derived from $\lambda$ vir phage.

$\lambda$ dv's 1 and 21 were the first to be mapped.  $\lambda$ dv1 is the original  $\lambda$ dv isolated by Matsubara and Kaiser (1968). Earlier reports showed that it exists as tandem dimers and higher oligomers (Champoux, 1970; Hobom . and Hogness, 1972). This is confirmed by our electron microscopy observation. In carrier cell 866, it exists predominantly as tandem dimers; no monomer has been found.  $\lambda$ dv21 exists mainly as monomers. In the preparations of both  $\lambda$ dv's, higher oligomers and interlocked molecules (catenated) are also seen. Native superhelical catenated molecules can usually be distinguished from oligomers because of the different super-helicities of the individual molecules. Catenated molecules can also be recognized in denatured preparations when some of the members are nicked and denatured, and some remain as closed circular forms. An example of catenated  $\lambda$ dv1 (and other  $\lambda$ dv's) is presented in Plate VI. The heteroduplex structures of  $\lambda$ dv's 1 and 21 with  $\lambda^+$  and  $\lambda$ b<sub>2</sub>b<sub>5</sub>c DNA's are straightforward as those depicted in Fig. 1 (a), (b) and (c). Some 1 (g) type heteroduplexes (without the nin deletion loop) have also been observed. Examples are shown in Plate II. Mapping results are presented in Table 4 .

### 2. $\lambda$ dv's derived from $\lambda$ nin<sup>-</sup>.

As stated in Methods for Mapping  $\lambda$ dv's, compound phage can be made by superinfecting  $\lambda$ dv carrier cells with deletion mutants of

TABLE 4  
Mapping Data for  $\lambda$ dv's

λdv*	Map Position ** on λ <sup>+</sup>			Total	Major Species;  Type of Inversion-Duplication ***	Size of Major Species		Hetero- genicity ***	Methods Of Mapping †	Single* Colony Purification
	Inverted λdv		Non-inverted λdv			Open Circles	Snap-back Linear (+ Unique Sequence, if Any)			
	Inverted Sequence	Unique Sequence								
1			73.7-88.6	14.9	tandem dimer (no monomer)			-	B-1, C	-
21			76.6-83.4	6.8	monomer	6.8		-	B-1	-
120			71.3-85.4 [14.1] gal = 7.8 at 84.8	21.9	partial dimer §	36.3		+	B-1, B-2	-
154			72.5-Δ-91.5	13.2	tandem dimer	26.1		+	B-2	+
203			69.3-83.8	14.5	tandem dimer			+	B-1	+
266			77.8-Δ-90.2 ΔΔ	6.6	monomer	6.5		+	B-1	+
104	73-Δ-92.6			13.8	Inverted dimer (5(a))	27.8	13.9	-	B-2	-
150	76.3-Δ-91.2 [9.1]	91.2-91.7 [0.5]		9.6	tandem diamphimer (5(dd))	37		+	B-2	+
161	72.9-Δ-90.3 [11.6]	90.3-90.7 [0.4]		12	monoamphimer (5(d))		12.4	+	B-2	+
249	78.3-83.1 [4.8]	72.6-78.3 [5.7]		10.5	Inverted diamphimer (5(bb), 5(bb'))	30.6	15.3	+	A	-
261 §§	70-Δ-91 [15.2]	91-91.2 [0.2]		15.4	monoamphimer (5(d))	31.5	15.4	+	B-2	+
280	73-Δ-89.7 [10.9]	89.7-90.4 [0.7]		11.6	monoamphimer (5(d))	23.1	11.6	+	A, B-2	+
292	73.7-Δ-91.4 [11.9]	91.4-91.8 [0.4]		12.3	monoamphimer (5(d))	24.8	12.4	+	A, B-2, C	-
200	75-Δ-91.7 [10.9]	91.7-92.1 [0.4]		11.3	monoamphimer (5(d))	23	11.3	+	A, B-2	+
204 LC	78-Δ-89.8 [5]	72.9-78 [5.1] 89.8-92 [2.2]		13.3	monoamphimer (5(c))	19.1		+	A, B-2	+
204 DK			75.8-Δ-90	8.4	tandem trimer	24.5		+	B-1	+
309	63.4-83.4 [20.0]	83.4-Δ-89.8 [0.6]		20.6	monoamphimer (5(d))	40.3	20.5	+	A, B-2	+

## Legend to Table 4

\* The genetic information of the  $\lambda$ dv's is listed in Table 3.  $\lambda$ dv's 1, 21, 120, 104, 249 and 292 were prepared from cells grown directly from stabs (-, last column of the table). All others were prepared from cells grown from tested single clones (+, last column of the table) as follows: Carrier cells were streaked on trypton agar plates; single colonies were picked with sterilized toothpicks and streaked against  $\lambda$  p gal 8 vir cII nin phage, except those carrying  $\lambda$ dv's 200, 203, 204 LC, 204 DK, on 4% "McConkey agar base" supplemented with 1% galactose.  $\lambda$ dv carriers survived the infection and were transduced to gal<sup>+</sup> and turned red after 24-36 hr at 37° C. Cells cured of  $\lambda$ dv were lysed. Resistant cells turned white. A  $\lambda$ dv positive clone from the master trypton agar plate was used to inoculate a culture for  $\lambda$ dv preparation. Cells carrying  $\lambda$ dv's 200, 203, 204 LC and 204 DK were purified similarly except that they were streaked against  $\lambda$  gal 8 cI sus 34 nin.

\*\* The formamide spreading technique was used throughout the mapping.  $\phi$ X174 and  $\phi$ X174 RF II were mounted on the same grids as internal standards and were taken as 11.2% of wild type  $\lambda$  DNA ( $\lambda$ <sup>+</sup>). The internal reference points used were the right boundary of i<sup>21</sup> (79.6) and the nin deletion, 83.8-89.6 determined by heteroduplexes of  $\lambda$ <sup>+</sup> and  $\lambda$ b<sub>2</sub>b<sub>5</sub>c and  $\lambda$  p gal 8 vir cII nin DNAs as compared to 79.8 and 83.8 to 89.2 by Davidson and Szybalski (1971). Each number represents an average of usually 20 to 80 measurements (and in no case less than 10). The standard deviations for the right boundaries

## Legend to Table 4 (continued)

of the  $\lambda$ dv's are 0.1 to 0.2%, and those for the left are larger, yet usually less than 0.5%.  $\Delta$  represents the nin deletion (83.8 - 89.6); numbers in brackets indicate segment lengths.

\*\*\* All  $\lambda$ dv's exist as monomers, dimers, trimers, higher oligomers and interlocked molecules except dv 1 which has no monomeric form. For details on types of inversion-duplication, see Fig. 5 and Results and Discussion; for heterogeneity, see Further Discussion.



§ The major species of  $\lambda$ dv 120 consists of two sets of  $\lambda$  gene complement and 1 set of gal genes. Molecules with 1 or 2 sets each of these genes and 1 set of  $\lambda$  genes only were seen at very low frequency.

§§ There is an unknown episome in  $\lambda$ dv 261 carriers. It has a size of about 140% of  $\lambda^+$  with no homology with either  $\lambda$  or the episome FDB (Sharp *et al.*, 1972). Its GC content is lower than that of  $\lambda$  as it denatures much more extensively than  $\lambda$ dv 261 when mounted in high formamide condition for denaturation map.

†cf. Methods for Mapping, Fig. 1, and Results and Discussion.

$\Delta\Delta$  Same as in Table 3.

Plate VI Interlocked  $\lambda$ dv molecules.

- A Interlocked closed circular  $\lambda$ dv1 molecules. The total length is that of a dodecamer. They are recognized as interlocked molecules because of the different superhelicity of the individual member.
- B  $\lambda$ dv 203. A closed circular molecule interlocks a single-stranded circular molecule. The preparation has been subjected to nicking, denaturation, and reneutralization treatments in which only one of the two interlocked molecules is nicked and denatured.
- C  $\lambda$ dv 309. A closed circular monomer interlocks a snap-back linear dimer at the w loop. A schematic representation of the linear dimer is seen as the top molecule in Fig. 5 (dd)'. 
- D  $\lambda$ dv 161. A snap-back dimer interlocks a snap-back monomer at the w loop. For a schematic representation of the monomer, see Fig. 5 (d'). The dimer probably has a structure as the bottom molecule of Fig. 5 (dd)'. with two collapsed w loops at the ends.  points to the w loop.

All molecules were mounted by formamide technique.



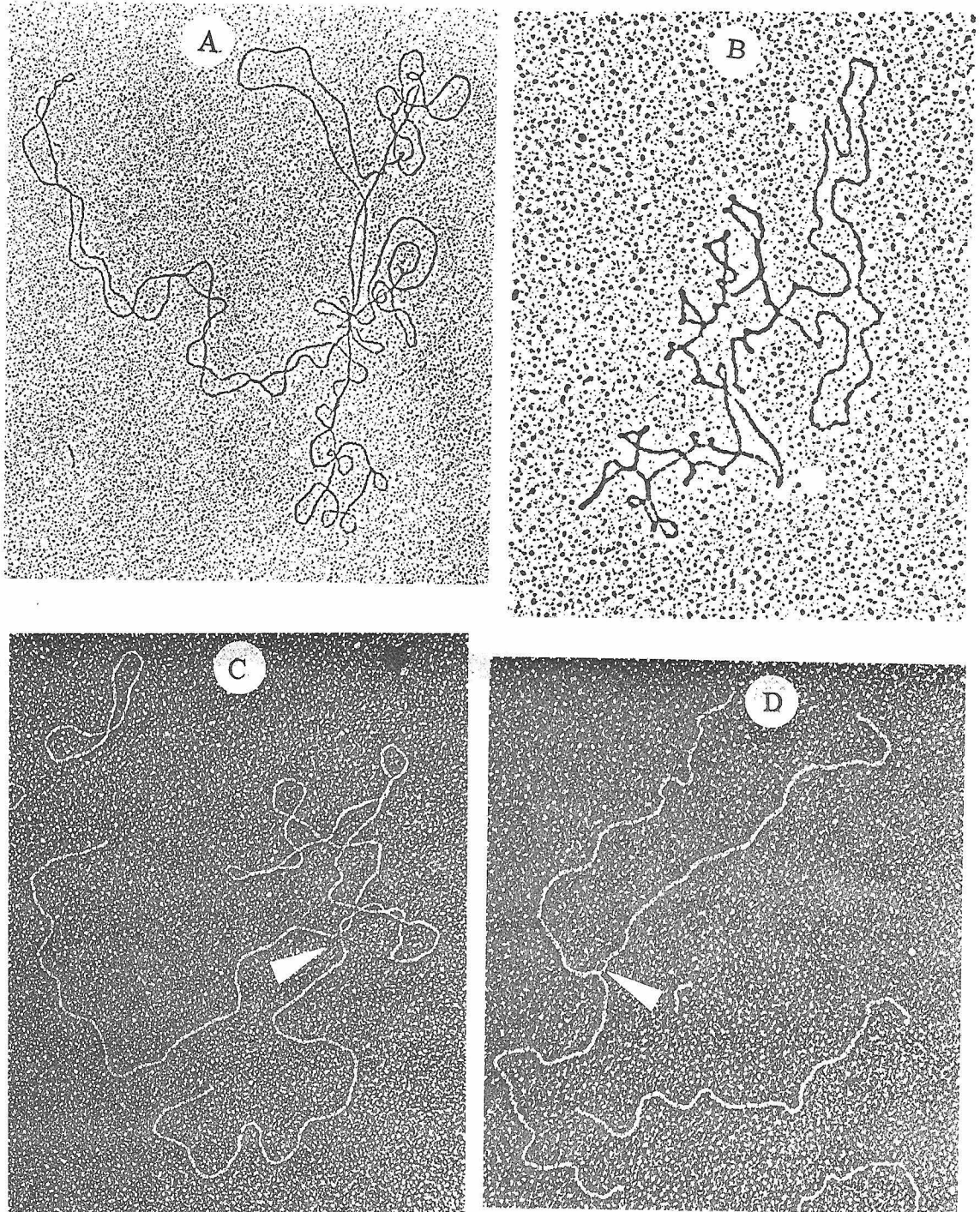


Plate VI

$\lambda i^{21}$ . The maximum amount of DNA that can be packed into a phage head is 106% of wild type  $\lambda$ . Therefore, the size of the smallest deletion in the  $\lambda i^{21}$  needed to permit the formation of a compound phage (a duplicate phage, Fig. 3) reflects the physical size of the  $\lambda dv$  monomer. Table 5 (Berg, personal communication) lists the number of  $\lambda dv$ 's derived from  $\lambda nin^+$  and  $\lambda nin^-$  phages that fit into  $\lambda i^{21}$ 's of various sizes. It may be seen that many  $\lambda dv$ 's derived from  $\lambda nin^+$  can fit into a phage with a 5% deletion, whereas those from  $\lambda nin^-$  frequently require phages with a 21% deletion, which is larger than one would expect from the genetic results (Tables 3, 5). In other words, many  $\lambda dv$ 's from  $\lambda nin^-$  phage were physically bigger than they are genetically. Table 6 (Berg, personal communication) presents the results of compound phage formation of a number of  $\lambda dv$ 's. The incongruity between the physical size and the genetic size of the  $nin^-$   $\lambda dv$ 's was resolved by our electron microscopic observations (to be described) that these  $\lambda dv$ 's contain inverted repeats. That is, they carry a duplication of some of all of the sequences in an inverted order so that one strand carries some sequences and their complements in an inverted order.

$\lambda dv104$ , which seemed to have a larger physical size than its genetic size in a compound phage test was studied after  $\lambda dv$ 's 1 and 21. It was isolated, sized and nicked as described. Quite unexpectedly, when hybridization preparations with  $\lambda b_2b_5c$  and  $\lambda c_{26}$  DNAs were examined in the electron microscope, hardly any heteroduplexes as depicted in Fig. 1 were found. To determine the cause of the problem,  $\lambda dv104$  along every step of the hybridization procedure was



TABLE 5

Minimum Deletion Size for Compound Phage Formation

(Dr. D. E. Berg, Personal Communication)

$\lambda$ dv Origin	Deletion				Number Analyzed
	5%	8%	11%	21%	
$\text{nin}^+$	15	8	0	0	23
$\text{nin}^-$	3	5	4	34	46

The superinfecting phages  $\lambda \underline{i}^{21}$ ,  $\lambda \underline{b}_{515} \underline{i}^{21}$ ,  $\lambda \underline{b}_{519} \underline{i}^{21}$  and  $\lambda \underline{b}_{538} \underline{i}^{21}$  with 5, 8, 11, and 21 % of  $\lambda^+$  DNA deleted, respectively. All four were  $\text{sus } \underline{P}$ . The compound phages were detected for their being  $\underline{P}^+$  and heterozygous for  $\underline{i}^{21} / \underline{i}^{\lambda}$ .

TABLE 6

% of  $P^+$  Compound Phage Carrying One or More  
Copies of  $\lambda dv$  Complement

(Dr. D. E. Berg, Personal Communication)

$\lambda dv$	Parental Phage*	Deletion Size of Superinfecting $\lambda i^{21}$ Phage			
		5%	8%	11%	21%
1	<u>nin</u> <sup>+</sup>	0	3	38	68
21	<u>nin</u> <sup>+</sup>	49	57	70	89
104	<u>nin</u> <sup>-</sup>	0	0	0	0**
161	<u>nin</u> <sup>-</sup>	0	0	0	17
200	<u>nin</u> <sup>-</sup>	0	0	0	12
204 LC	<u>nin</u> <sup>-</sup>	0	0	11	73
204 LC	<u>nin</u> <sup>-</sup>	0	0	15	75
266	<u>nin</u> <sup>-</sup>	39	68	74	89
309	<u>nin</u> <sup>-</sup>	0	0	0	0

\* cf . Table 3 for detailed genotype of parental phages.

\*\* About 0.1 % of the  $P^+$  plaque was compound phage . However,  
the plaques were strongly mottled with more than 99% segregants.

examined. The nicked DNA preparation contained both open circles and un-nicked closed circles (Plate VII B). However, denaturation by alkali followed by reneutralization of the nicked DNA preparations gave mainly un-nicked closed circles and double-stranded linear molecules half the size of the major open circles (Plate VII A).

There were also a few short single-stranded fragments when mounted by the formamide technique. Some of the double-stranded molecules had a single-stranded gap at one end or in the middle (Fig. 8). No single-stranded circles or long linear single strands were seen.

When the same preparation was mounted by aqueous technique, only closed circles, double-stranded linears with or without single-strand bushes at one end or in the middle were seen. The same result was obtained when self renaturation of the preparation was carried out. The interpretation of these observations is that  $\lambda$ dv104 molecule contains an inverted repeat. The smallest unit is a head-to-head inverted dimer; i.e. it contains complementary sequences in an inverted order on the same strand. After nicking, denaturation and reneutralization, the complementary single strand "snaps back" on itself to form a double-stranded linear molecule of half the size of the original open circles (Figs. 5 and 8). The observation that in the denatured preparation there were no double-stranded "snap-back" linears with two single-stranded ends, is consistent with this interpretation.

To prove this hypothesis, electron microscopy mounting conditions were sought under which the double-stranded linear molecules that had snapped back were denatured to give single-stranded circular

Plate VII  $\lambda$ dv 104

A Nicked, denatured, and reneutralized preparation.

B Nicked, preparation.

A and B were mounted from 40 % onto 10 % formamide.

C Nicked, denatured, and reneutralized preparation. mounted from 90 % onto 55 % formamide at 40° C, showing one fully denatured, single-stranded circle and one undenatured snap-back linear.

D The same as C. The small circle is  $\phi$ x 174.

E Nicked, denatured, and reneutralized preparation mounted from 85 % onto 50 % formamide at 40 °C, showing a partially denatured snap-back molecule. The small circle is  $\phi$ x 174.

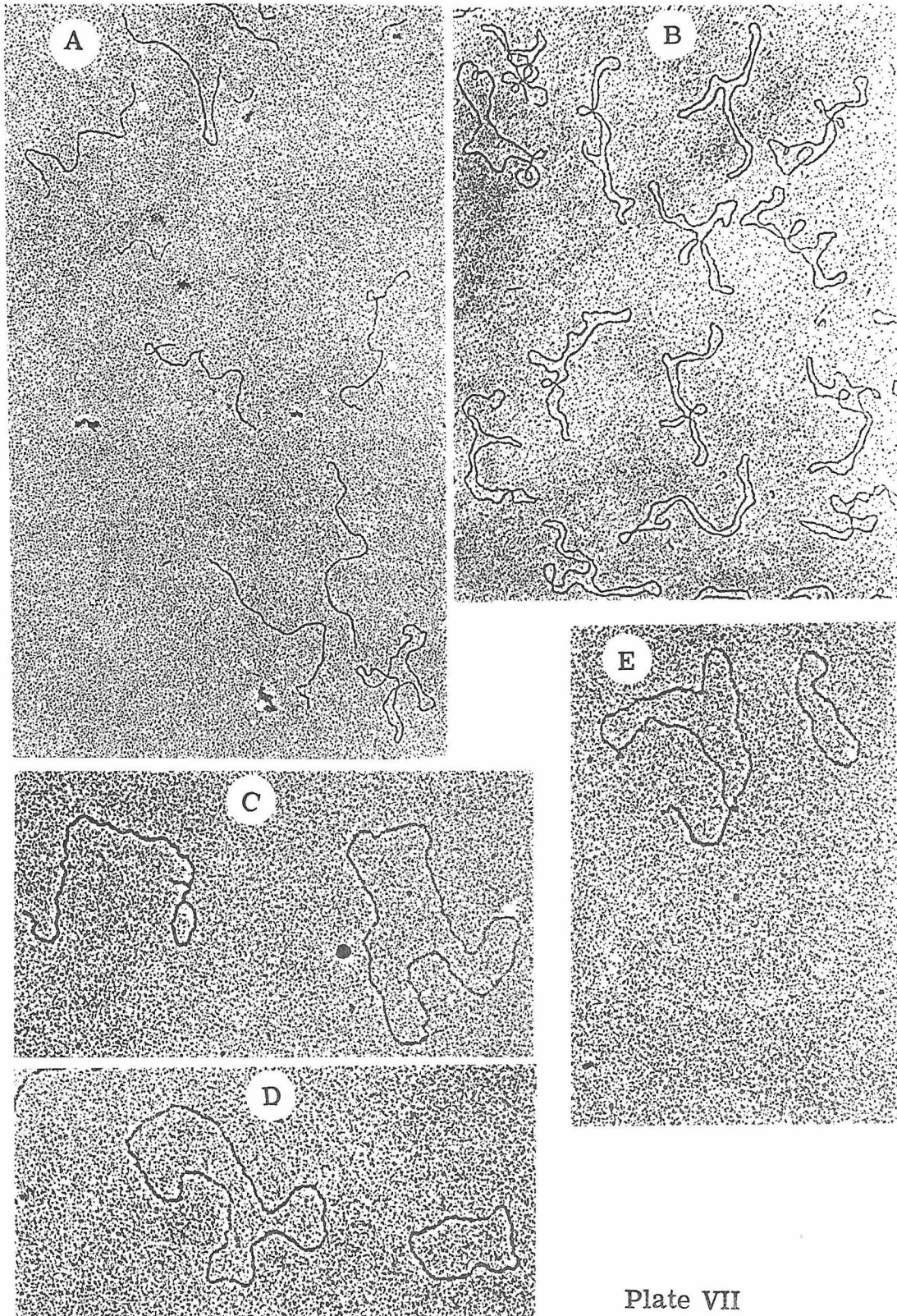


Plate VII

and linear molecules equivalent in length to that of a double-stranded open circle. Such conditions were found: Partial denaturation was achieved by spreading from 85% onto 50% formamide at 40° C with the salt concentrations described in Experimental Procedures. Total denaturation (for a fraction of the molecules) was observed by spreading from 90% onto 55% formamide at 40° C. Examples are presented in Plate VII C, D and E.

Careful examination of the denatured  $\lambda$ dv104 mounted from 40% onto 10% formamide indicates that it is completely inverted, with no unique sequence (no single-stranded loop) at either end (Fig. 5 (a)). Therefore,  $\lambda$ dv104 is an inverted dimer.

Thus, our observation fully explains the novel behavior of  $\lambda$ dv104 in the compound phage test. "Addition recombination" between  $\lambda$  i<sup>21</sup> and  $\lambda$ dv DNA *always* leads to the insertion of the whole inverted dimer containing two sets of the gene complement. Under no circumstances can one set of gene complements be inserted into the  $\lambda$  DNA as the non-inverted  $\lambda$ dv's 1 and 21 can. Therefore, the physical size of  $\lambda$ dv104 is larger than its genetic size.

After the discovery of the inversion duplication in  $\lambda$ dv104, other  $\lambda$ dv preparations were routinely screened for inverted repeats by electron microscopic examination of the denatured molecules.

Out of a total of 14  $\lambda$ dv's derived from  $\lambda$  nin<sup>-</sup> phage examined, ten turned out to contain inverted repeats. There are four types of inversions (Table 4): (a) complete inversion and duplication (Fig. 5 (a)), (b) partial inversion and duplication with a long non-inverted, non-duplicated (or unique) sequence (5.7% of  $\lambda$ <sup>+</sup>) on the

left covering the  $\lambda$  immunity region (oriented according to the  $\lambda$  map) (Fig. 5(b)), (c) partial inversion and duplication with a long (5.1%) and a short (2.2%) unique sequence on the left and the right, respectively (Fig. 5 (c)), (d) partial inversion and duplication with a very short ( $<0.7\%$ ) unique sequence on the right hand side of the  $\lambda$ dv close to the nin deletion (Fig. 5 (d)). Figure 5 illustrates the four types of structures bearing inverted repeats that we have encountered, and also the structures observed by snap-back of a single strand from monomeric and dimeric forms of these  $\lambda$ dv's. Examples are presented in Plates VI C, D, VII A, VIII, IX C, D, E, F and XV.

I would like to introduce, at this point, the word "amphimer" to describe plasmids containing inverted partial repeats. I use "amphimer" to stress the facts that the smallest unit of such a plasmid has the characteristics of both a monomer and a dimer, and that the same strand contains sequences derived from the plus strand and the minus strand of the parental chromosome. Therefore, types 5 (b), (c) and (d)  $\lambda$ dv's are monoamphimers whereas type 5 (a) is an inverted dimer.

Type 5 (d) amphimers have such short unique sequences at the right end that they were first mistaken for type 5 (a) until they were hybridized to  $\lambda$  DNA as described later.

Types 5 (b) and (c) ( $\lambda$ dv249 and 204 LC, respectively) were mapped directly by hybridizing to wild type  $\lambda$  (mapping method B-1 (a)) or their parental phage (method A). The heteroduplexes between

Fig. 5 Types of  $\lambda$ dv's with inverted repeats and the structures of snap-back molecules observed after nicking, denaturation, and reneutralization.

The definition for amphimer is given in the text.

(a) to (d) and (a') to (d') are the native and snap-back forms of the inverted dimer and the monomeric amphimers.

(aa) to (dd) , and (aa)' to (dd)' are the tandem dimeric forms of (a) to (d), and (a') to (d') respectively .

(bb<sub>1</sub>) and (bb<sub>2</sub>) are two native inverted dimers of (b).

(bb<sub>1</sub>)' and (bb<sub>2</sub>)' are the snap-back forms of (bb<sub>1</sub>) and (bb<sub>2</sub>).

The DNA sequences are labeled with letters x, y, z, etc. ; x' is the complement of x, and so on. Only one strand is lettered. The same lettering among the different types of  $\lambda$ dv's does not necessarily represent identical sequence.



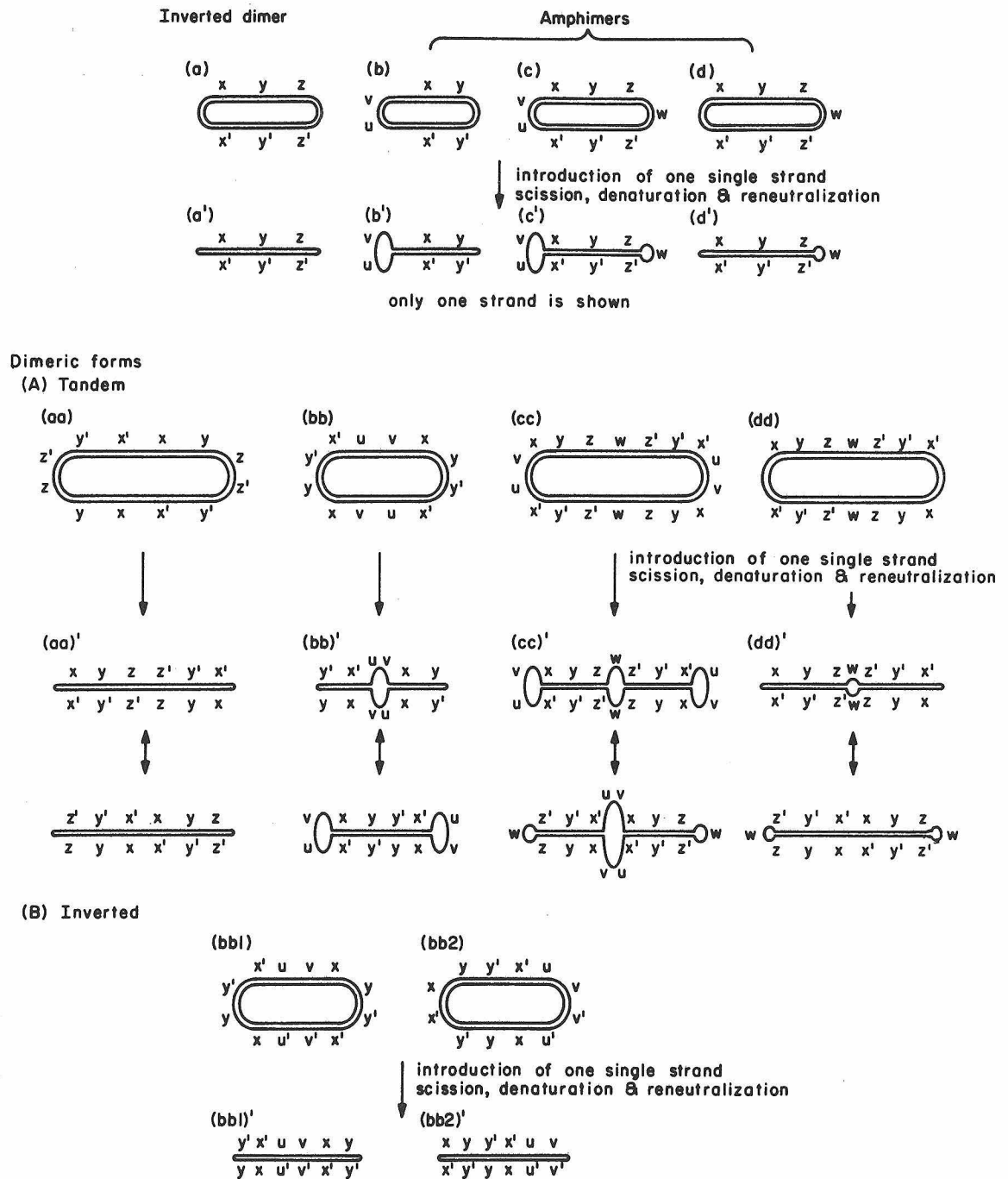


Fig. 5

Plate VIII Denatured  $\lambda$ dv 204 LC.

A From a hybridization preparation with  $\lambda$  DNA. The small circles are  $\phi$  x 174 and  $\phi$  x 174 RF II, The long linears are  $\lambda$  DNA.

B, C, and D From an isolated  $\lambda$ dv preparation. (c') and (cc)' refer to Fig. 5 (c') and (cc)' respectively. Molecule I is a branch migrated intermediate of the two (cc)' forms. Molecules II and III are generated when I is nicked at positions 1, 2 and 1, 3 respectively, and subsequently denatured, and reneutralized.

▲ points to the single stranded region in molecules II and III. All molecules were mounted from 40 % onto 10 % formamide.

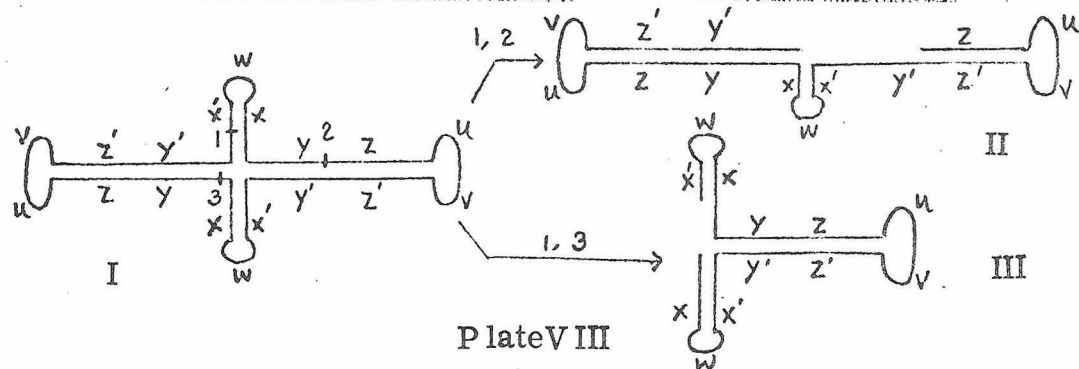
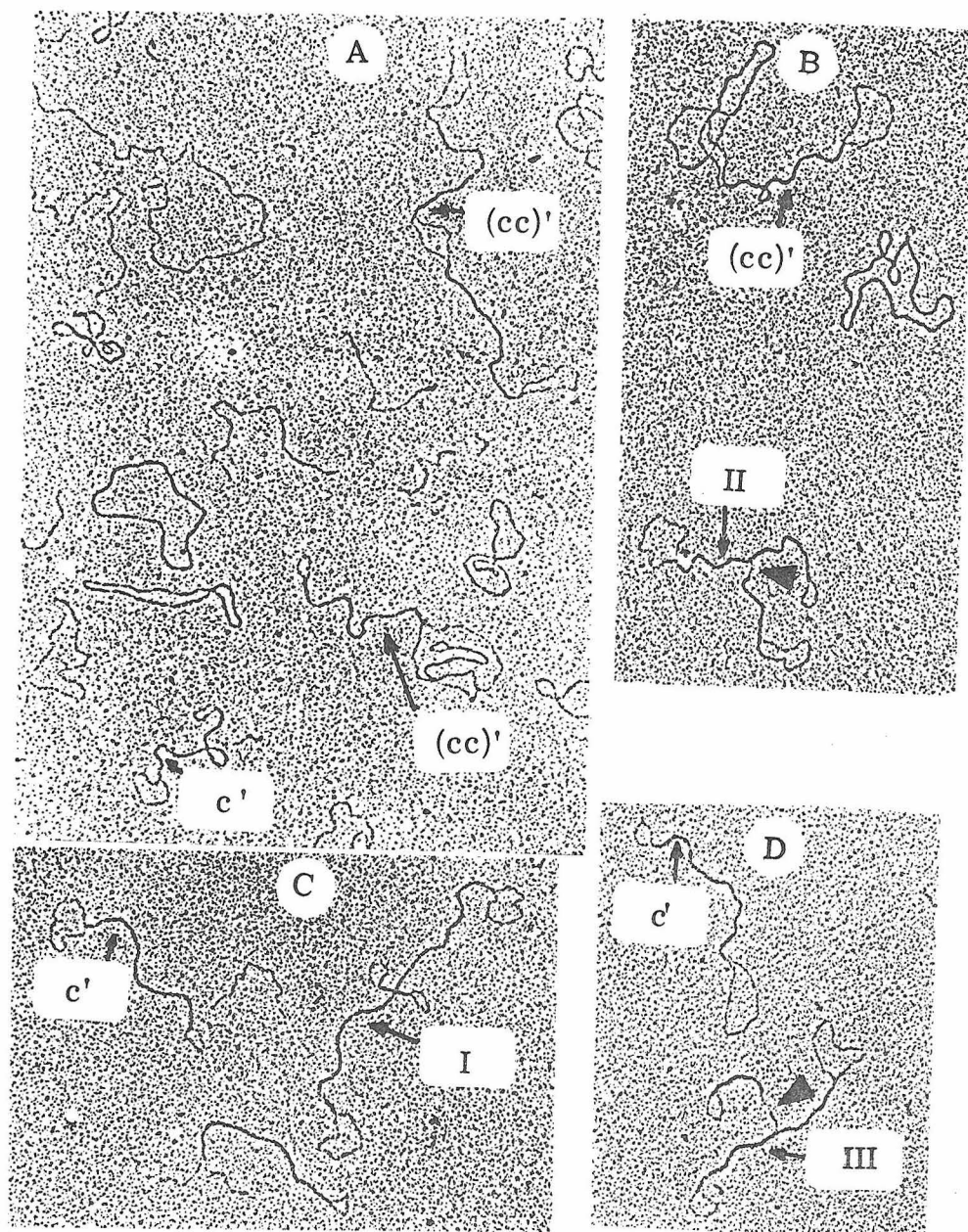


Plate V III

$\lambda$ dv monoamphimers and  $\lambda$  DNAs seen are depicted in Fig. 6.

Branched migrated forms as depicted in Fig. 6 (a) and 9 (c) are often seen. For type 5 (c)  $\lambda$ dv,  $(100 - r_1)$  and  $(100 - r_2 - \underline{uv})$  represent the right and left end points, respectively; duplex  $\underline{uvxyzw}$  is the genetic size; duplex  $\underline{uvxyzw}$  plus single strand  $\underline{xyz}$  is the physical size of the monoamphimer which is equal to the double-stranded open circular molecule; duplex loops  $\underline{uv}$  and  $\underline{w}$  are the non-inverted, non-duplicated sequences. The duplex loops  $\underline{uv}$  and  $\underline{w}$  can be measured more accurately than the single-stranded loops  $\underline{uv}$  and  $\underline{w}$ . The latter have a tendency to collapse partially and to be variable and short in length. Data analysis for type 5 (b) is similar. The results are presented in Table 4. Examples of heteroduplexes of mono- and diamphimers with  $\lambda$  DNAs are shown in Plates IX A, B, and X.

$\lambda$ dv249 is extraordinary. The open circle measures 30.6% of  $\lambda^+$ . Denatured preparation contains 85 to 90% double-stranded linears of 15.3% of  $\lambda^+$  and 15 to 10% of denatured mono- and tandem diamphimers (Fig. 5 (b') and (bb')). The monoamphimer has a linear sequence of  $\underline{xyy'x'uv}$  ( $\underline{x'}$  and  $\underline{y'}$  represent the complementary sequences of  $\underline{x}$  and  $\underline{y}$ ) and a length of 10.5% of  $\lambda^+$  ( $\underline{uv} + \underline{xy}$ , where  $\underline{uv}$  is 5.7% and  $\underline{xy}$  is 4.8% of  $\lambda^+$ ). The tandem diamphimer has a sequence of  $\underline{xyy'x'uvxyy'x'uv}$  and a length of 15.3% of  $\lambda^+$  ( $2 \times \underline{xy} + \underline{uv}$ ). Hybridization with  $\lambda^+$  gives only one mapping result (Table 4). Therefore, we conclude that the majority of  $\lambda$ dv249 is an inverted diamphimer with a linear sequence of either  $\underline{y'x'uvxyy'x'v'u'xy}$  or  $\underline{xyy'x'uvv'u'xyy'x'}$  (Fig. 5 (bb<sub>1</sub>)' or (bb<sub>2</sub>)').

Fig. 6 Heteroduplexes of types ( Fig. ) 5 (b) and (c)  $\lambda$ dv's with  $\lambda$  DNA's.

- (a)  $\lambda$ dv 249 ( type 5 (b) ) /  $\lambda c_{26}$  (wild type) or parental phage ( $\lambda p$  gal 8 vir cII nin) DNA. The single strand  $r_3$  in the heteroduplex with  $\lambda c_{26}$  is longer than that in the heteroduplex with  $\lambda nin^-$  by the length of the nin deletion .
- (a') A branch migrated form of (a).
- (b) The same as (a) with completed branch migration.  $\lambda$ dv 249 does not extend beyond the nin deletion, therefore, there is no nin deletion loop in the heteroduplex.
- (c) and (d)  $\lambda$ dv 204LC ( type 5(c) ) /  $\lambda c_{26}$  or parental phage (  $\lambda p$  gal 8 cI 857 nin ) DNA. The single strand  $r_1$  is the same in heteroduplexes with  $\lambda c_{26}$  and  $\lambda nin^-$ .  $r_2$  in heteroduplex with  $\lambda c_{26}$  is longer by the length of the nin deletion. Branch migrated forms similar to (a') and Fig. 9(c) have also been observed.
- (e)  $\lambda$ dv204LC / parental phage DNA with completed branch migration.
- (f)  $\lambda$ dv 204LC /  $\lambda c_{26}$  with completed branch migration.

Notation is the same as in Fig. 1 and Fig. 5.

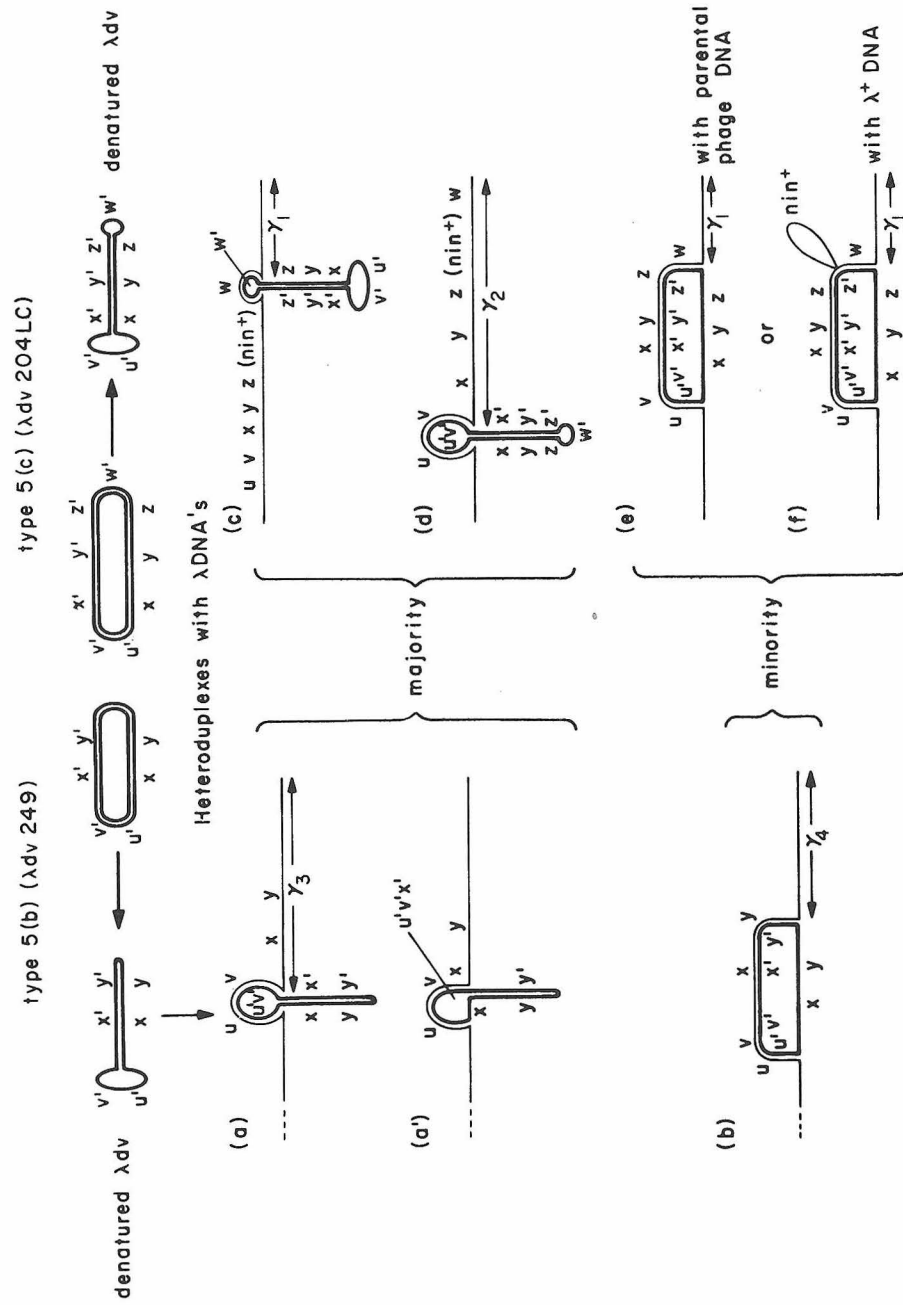


Fig. 6

The former has a mirror image plane between  $\underline{y} - \underline{y}'$  and can be formed by recombination between a plus strand of one monoamphimer and a minus strand of another in the inverted, duplicated region. The latter has a mirror image plane between  $\underline{x} - \underline{x}'$  and  $\underline{v} - \underline{v}'$  and can be formed by an error in replication. I do not know which is the real sequence of the double-stranded linears observed. This conclusion is supported by the observation of trimers containing two monomeric sequences in tandem and the third one in an inverted order. It has a sequence of either  $\underline{y' x' u v x y y' x' v' u' x y y' x' v' u' x y}$  or  $\underline{x y y' x' u v v' u' x y y' x' v' u' x y y' x'}$ . The trimer, along with monomers, tandem and inverted dimers, are shown in Plate IX C, D, E. and F.

In principle, the compound phages of the inverted  $\lambda dv$ 's can be used for mapping the  $\lambda dv$ . Figure 4 (d) and (e) show heteroduplexes of compound phages containing types 5(a) and 5 (b)  $\lambda dv$  and  $\lambda i^{21}$  in which  $(100 - r)$  and  $(100 - r - D)$  or  $(100 - r - t)$  represent the right and left end points of the inverted  $\lambda dv$ 's. However, in practice it is not easy to make such compound phages, especially when the  $\lambda dv$  is big. For instance, attempts to make the stable duplicate phage  $\lambda b_{538} i^{21} - dv$  104 were not successful (Table 6). Furthermore, when compound phage of  $\lambda dv$ 's 249 or 292 was grown, three phage bands were visible in the CsCl density gradient. The bands were not well resolved and contained segregants, abnormal compound phages (compound phages with fractions of the  $\lambda dv$  DNA deleted),  $i^\lambda$  recombinants, and probably true compound phage as determined by electron

Plate IX Denatured  $\lambda$ dv 249 and heteroduplexes with  $\lambda$  DNA.

- A  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 249. The inset is an enlarged tracing of the heteroduplex. A schematic representation is seen in Fig. 6 (a').  $\rightarrow$  points to the junction of  $\lambda$  and  $\lambda$ dv DNA's. One of the single strands at the fork belongs to  $\lambda$ dv and is so labeled.
- B  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv249 . For a schematic representation, see Fig. 6 (a).  $\blacktriangle$  points to the junction of  $\lambda$  and  $\lambda$ dv DNA's.
- C, D, and E Denatured  $\lambda$ dv 249. (bb), (b') and (bb)' refer to Fig. 5 (bb) , (b') and the top molecule of (bb)' respectively. (bb), a close circular molecule, interlocks II, a snap-back inverted diamphimer with a sequence of either (bb<sub>1</sub>)' or (bb<sub>2</sub>)' of Fig. 5. III, a trimer with two monomeric sequences in tandem and the third one in an inverted order.
- F Two interlocked (bb)' molecules.

All molecules were mounted by formamide techniques.



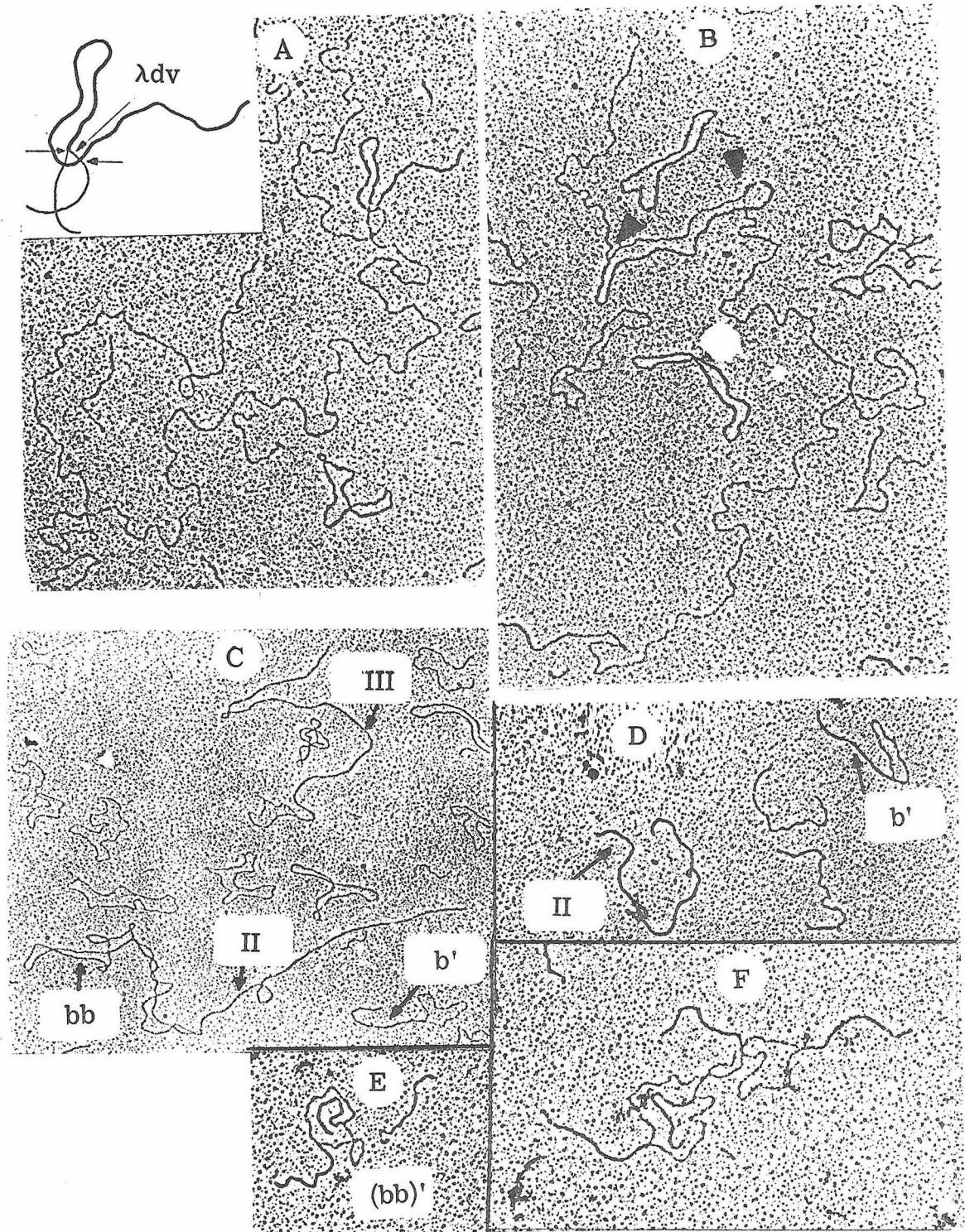


Plate IX

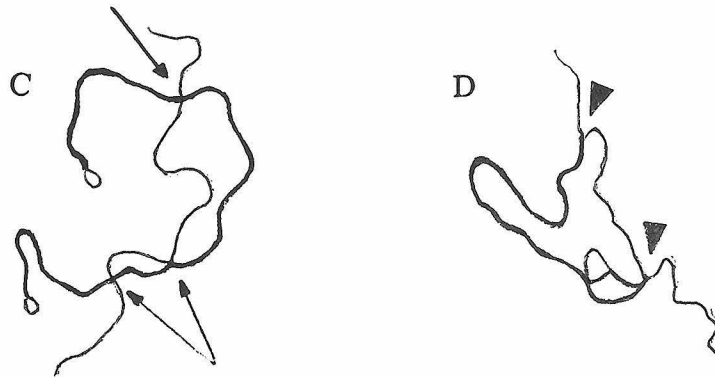


Plate X Heteroduplexes of  $\lambda c_{26}$  and type ( Fig. 5(c) )  $\lambda dv$   
(  $\lambda dv$  204 LC ) DNA's.

- A  $\lambda c_{26} / \lambda dv$  monoamphimer. For a schematic representation, see Fig. 6 (e).  $\blacktriangle$  points to the junction of  $\lambda$  and  $\lambda dv$  DNA's.
- B  $\lambda c_{26} / \lambda dv$  monoamphimer. A schematic representation is seen in Fig. 6 (c).  $\blacktriangle$  Points to the duplex  $w$  loop.
- C  $\lambda c_{26} / \lambda dv$  diamphimer ( in the bottom form of Fig. 5 (cc)' ).  
 $\rightarrow$  points to the junction of  $\lambda$  and  $\lambda dv$  DNA's. Branch migration may have occurred at the bottom junction as indicated in the enlarged, illustrative tracing above.
- D  $\lambda c_{26} / \lambda dv$  monoamphimer . This is a branch migrated intermediate of the two heteroduplexes depicted in Fig. 6(c) and (d).  
 An enlarged, illustrative tracing is also shown above.

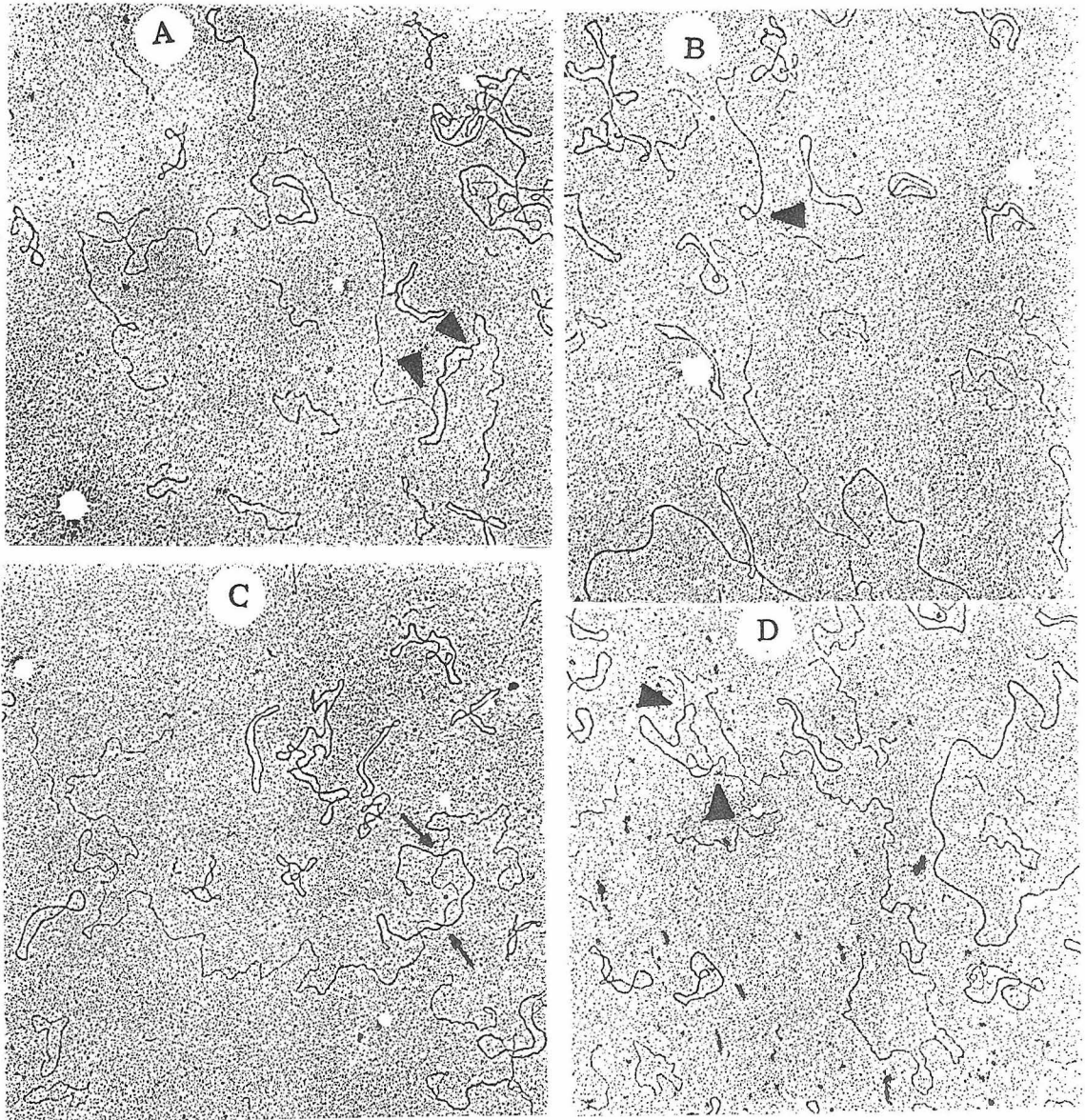


Plate X

microscopy. Examples of heteroduplexes of  $\lambda_{b_{538}}^{i^{21}} - \lambda_{dv}^{inverted}$  with  $\lambda_{b_2b_5c}$  are shown in Plate XI. Until stable compound phages can be made, it is not feasible to use this method to map  $\lambda_{dv}$  amphimers.

Another method for mapping type 5 (a) ( and type 5 (d) mistaken for (a))  $\lambda_{dv}$ 's is by comparing the partial denaturation pattern of the  $\lambda_{dv}$  (in either an open circular or a "snap-back" linear form) with that of  $\lambda$  DNA. Under certain highly denaturing conditions, DNA segments high in AT content are denatured to give single-stranded loops, whereas those high in GC content remain double-stranded. As the conditions become more and more denaturing, more segments are denatured. Therefore, under a certain denaturing condition,  $\lambda$  DNA shows a certain pattern of double-stranded segments and single-stranded loops (Inman, 1966; Schnös and Inman, 1970). A high formamide mounting condition to denature  $\lambda$  and  $\lambda_{dv}$  DNAs *in situ* was found and is described in Experimental Procedures. The partial denaturation maps of wild type  $\lambda$  ( $\lambda_{c_{26}}$ ) and  $\lambda$  p gal 8 vir cII nin DNAs are presented in Fig. 7. Typical molecules are shown in Plate XII. Although the general loop patterns are fairly reproducible, the accuracy (probably no better than 2%) of mapping  $\lambda_{dv}$  is not as high as desirable. Therefore, the heteroduplex methods were pursued. The logic is the following. Although completely inverted  $\lambda_{dv}$ 's which are nicked not more than once do not form heteroduplexes with  $\lambda$  DNA due to snapping back, more heavily nicked  $\lambda_{dv}$ 's are expected to do so. The sequence of events is presented in Fig. 8 (c). Data analyses for the heteroduplexes in Fig. 8 (c)

Plate XI Heteroduplexes of DNA's of  $\lambda$  and compound phage  $\lambda$   
containing  $\lambda$ dv with inverted repeats.

A  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 292 /  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$

B  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 249 /  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$  . For a partial schematic representation , see Fig. 4 (e).

C  $\lambda \underline{b}_{538} \underline{i}^{21} - \lambda$ dv 249 ( abnormal, with a shorter inverted duplication. ) /  $\lambda \underline{b}_2 \underline{b}_5 \underline{c}$  .

▲ point to the inserted  $\lambda$ dv. △ points to the small deletion loop formed due to the difference between deletions  $\underline{b}_{538}$  and  $\underline{b}_2$ . All molecules were mounted from 45 % onto 17 % formamide.



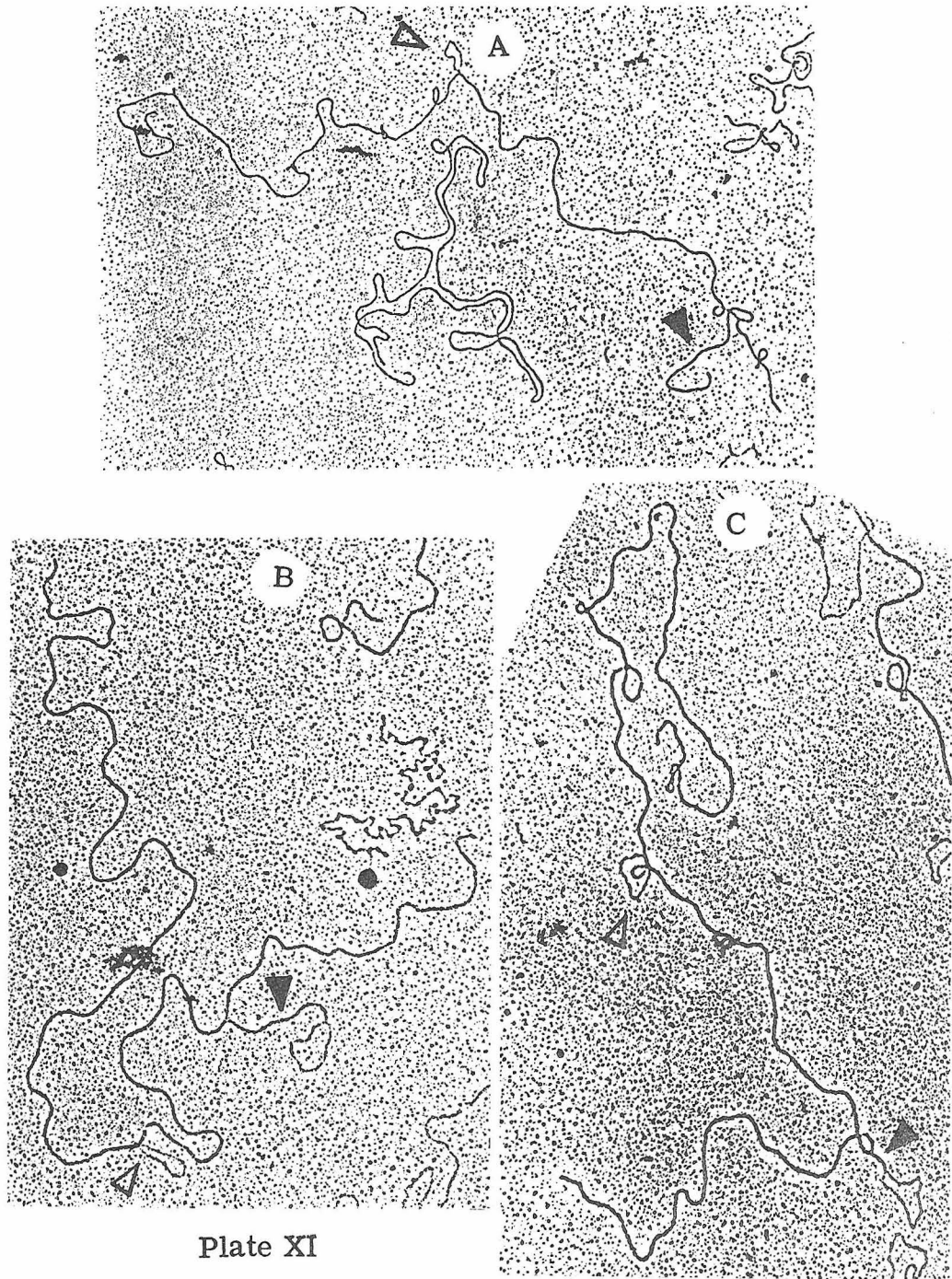


Plate XI

Fig. 7 Partial denaturation maps of  $\lambda^+$  (A) and  $\lambda$ p gal 8 vir cII nin (B).

Conditions for partial denaturation in situ are described in Experimental Procedures. Data are processed as follows:

The lengths of the duplex segments ( $D_i$ ) and the single-stranded loops ( $l_i$ ) were measured and their percentages were calculated as  $D_i \% = D_i / (\sum D_i + \sum l_i) 100\%$  and  $l_i \% = l_i / (\sum D_i + \sum l_i) 100 \%$ . The data were plotted as such for 32  $\lambda$ p gal 8 vir cII nin molecules ( the elevated horizontal lines represent the loop regions ). For  $\lambda$ c<sub>26</sub>, the data for 13 out of 29 molecules were normalized as follows:

All 29 molecules had a very short  $D_2$  segment ( 0.5 to 1 % ) and a short  $l_1$  segment. The junction ( J ) of  $D_2$  and  $l_2$  of many molecules were at 51 % as calculated above. Therefore, J at 51 % was used as an internal reference point. Normalization factors for duplex and single-stranded DNA (  $N_D$  and  $N_l$  respectively ) were taken as  $51 / ( D_1 + l_1 + D_2 )$  and  $\{ (100 - 51) - \sum_{i=2} D_i \times N_D \} / \sum_{i=2} l_i$  respectively. The normalized duplex segments (  $D_{iN} = D_i \times N_D$  ) and the normalized single-stranded loops (  $l_{iN} = l_i \times N_l$  ) ( except  $D_1$  and  $l_1$  that were used without normalization ) were used for plotting the map. The molecules were lined up. The histograms were obtained by taken one point every 0.5 % of the molecules.

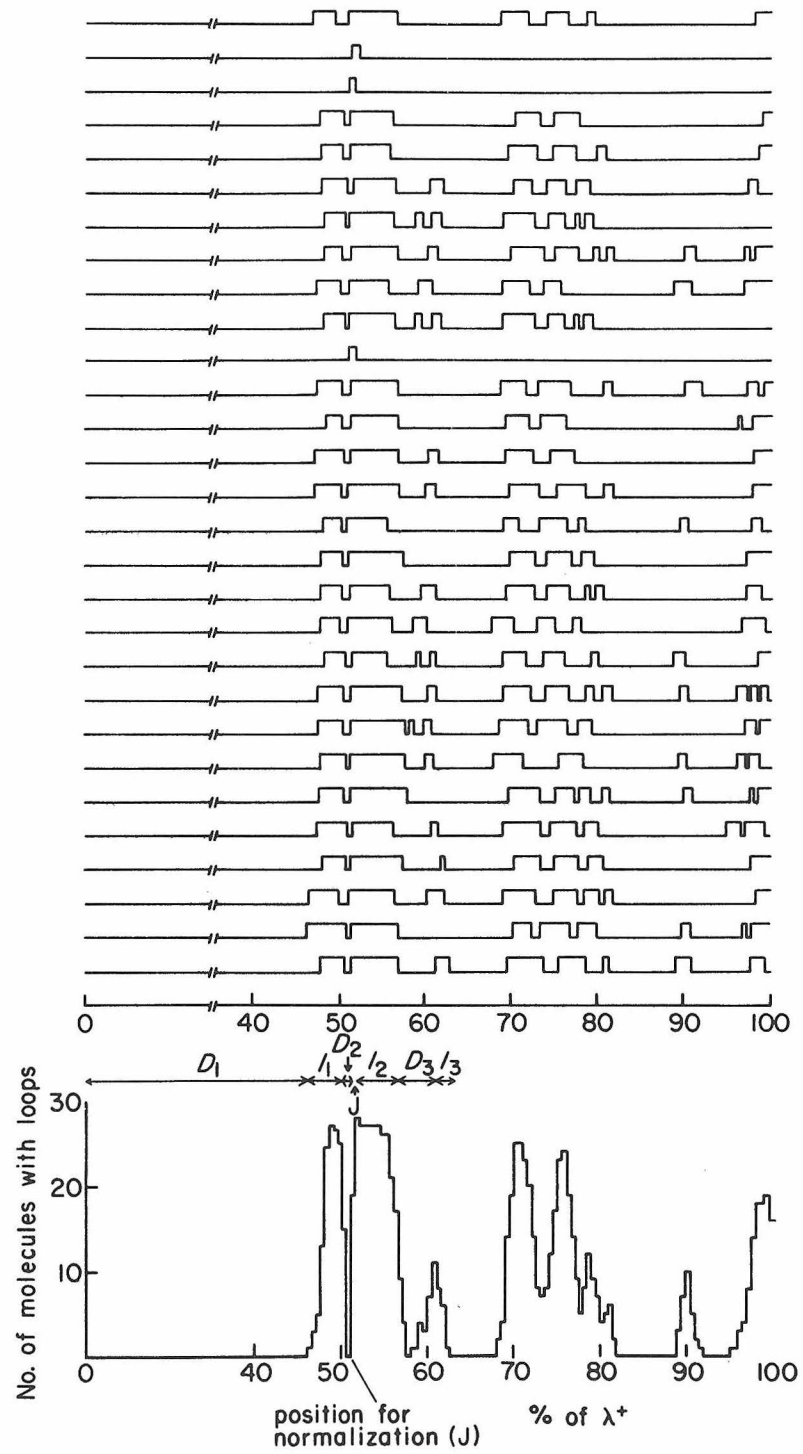


Fig. 7A



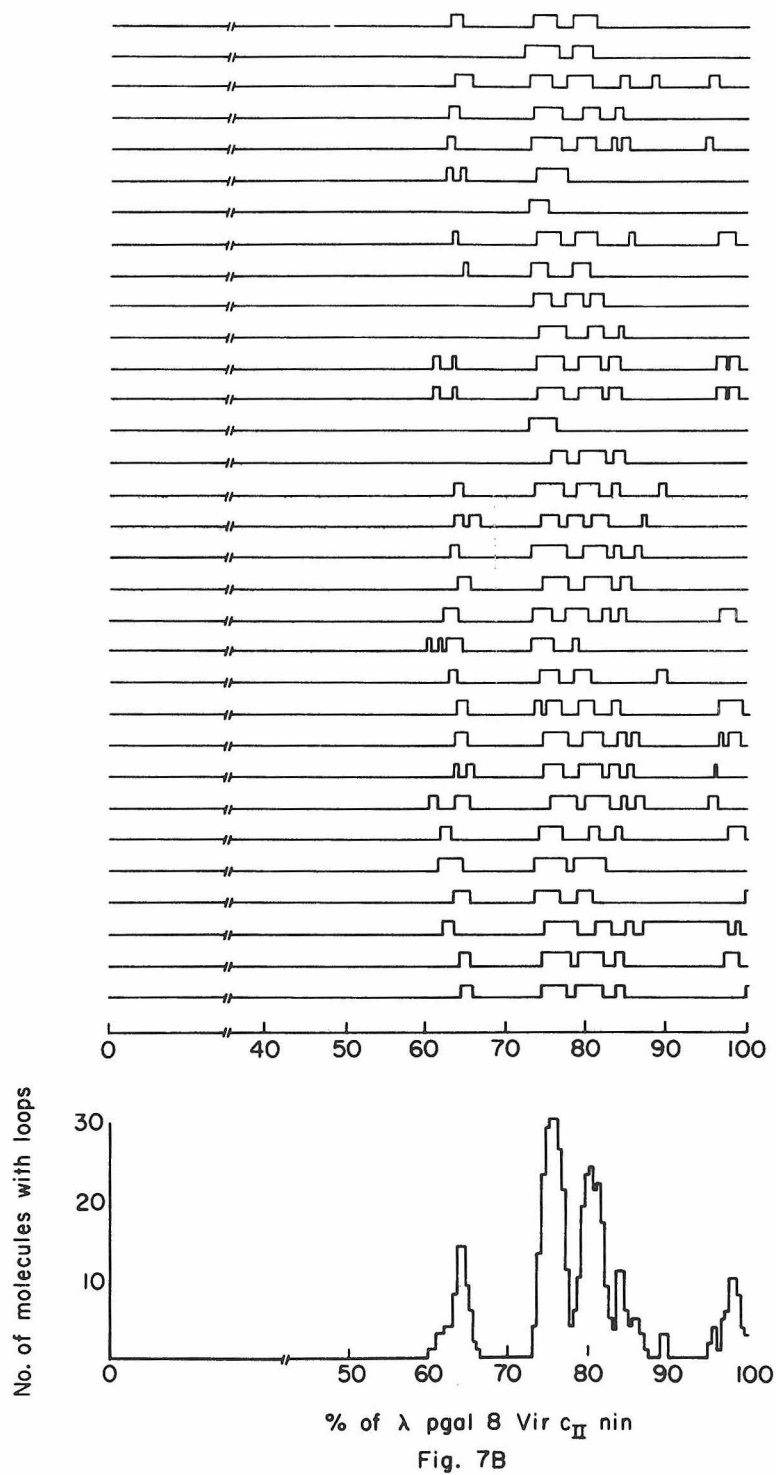



Plate XII Partially denatured  $\lambda$ DNA's.

A  $\lambda$ c<sub>26</sub> ( wild type DNA ).

B  $\lambda$ pgal 8 vir cII nin.

The mounting condition has been described in Experimental Procedures.  points to the right end of the molecule ( oriented according tot the  $\lambda$  map) Schematic representations are seen in Fig. 7 ( A ) and ( B ) respectively.

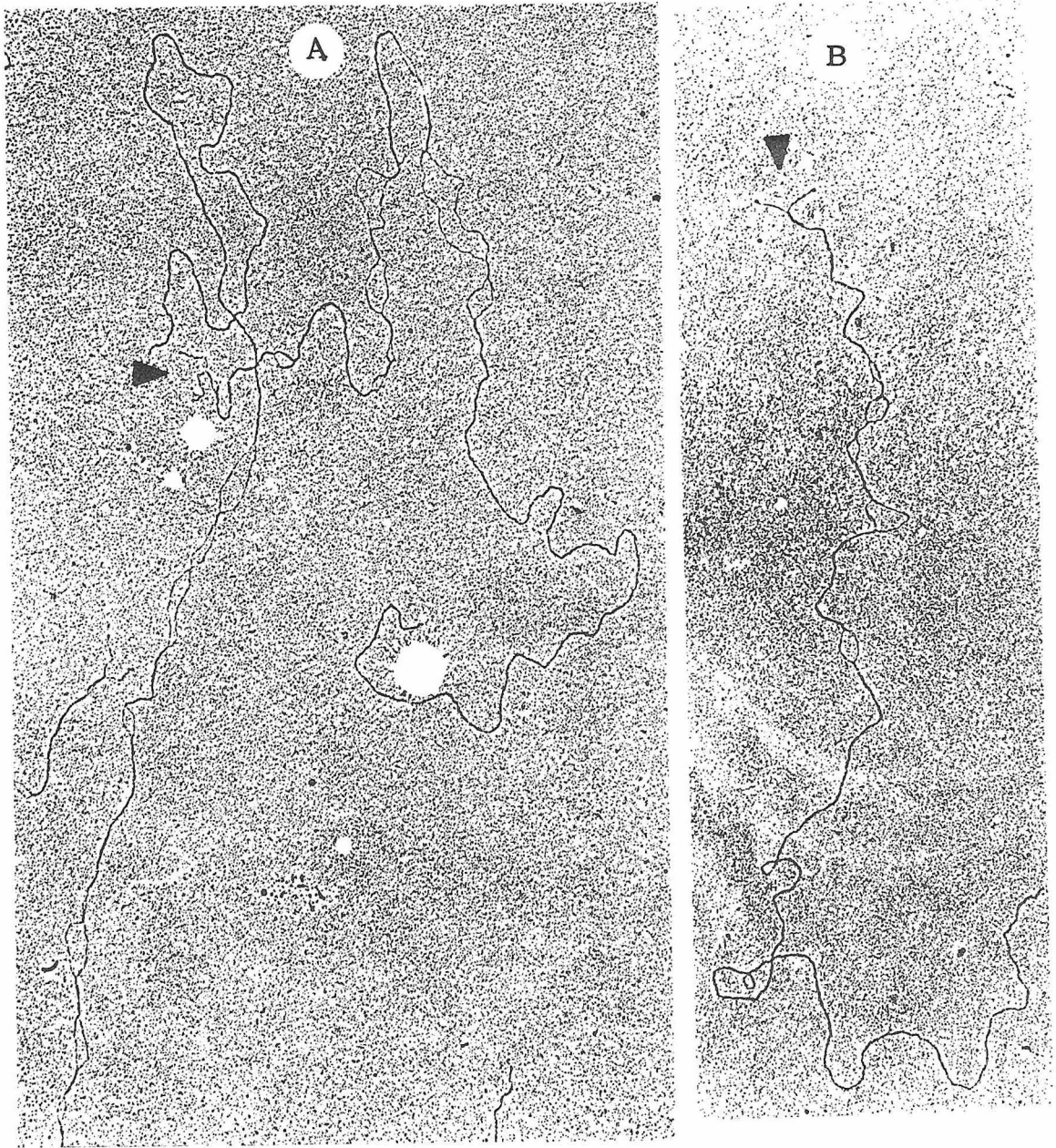


Plate XII

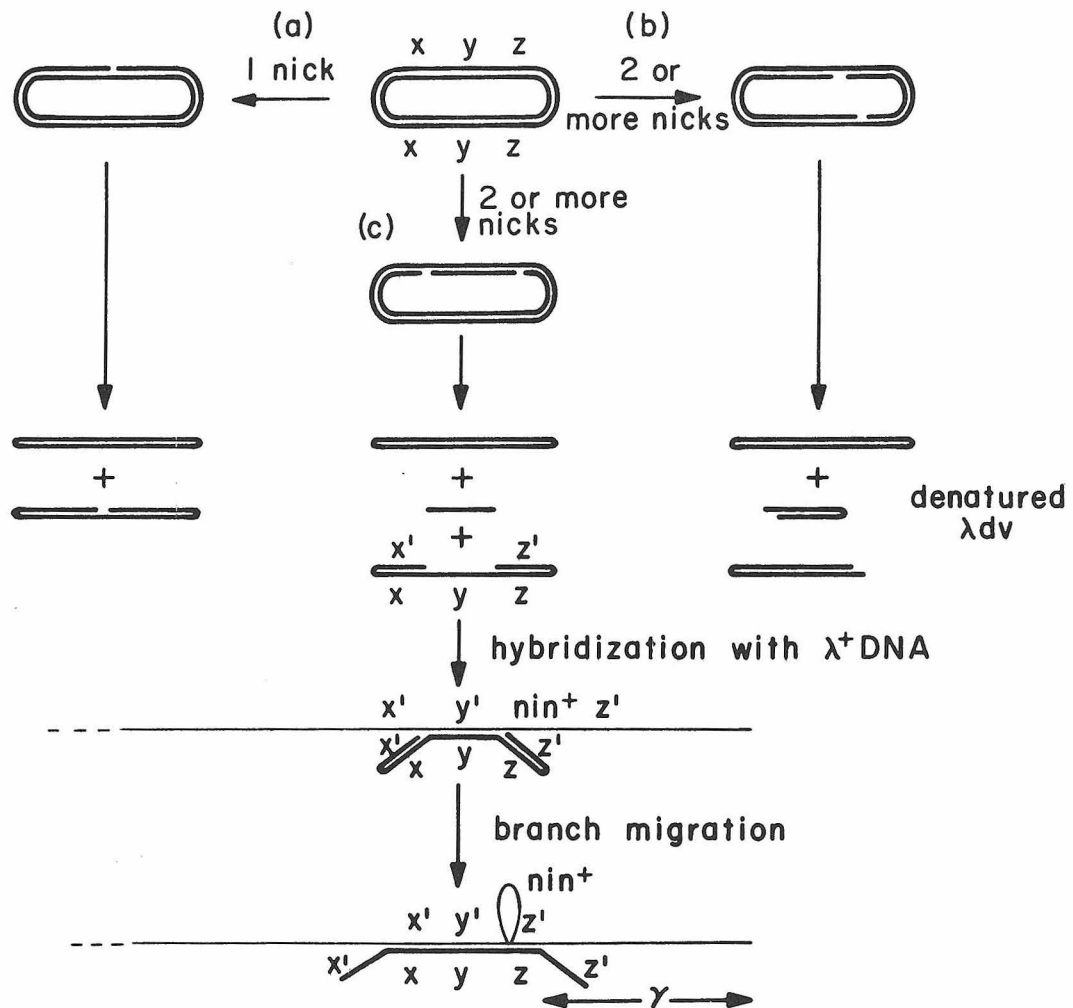


Fig. 8 Heteroduplex of  $\lambda$  and  $\lambda dv$  with complete inversion and duplication.

The notation is the same as in Fig. 5.

are the same as for non-inverted as well as for type 5 (b) and (c)  $\lambda$ dv's. Therefore, DNA preparations of type 5 (a) (and (d)) should be nicked more heavily. However, being overcautious, I actually introduced, on the average, only one nick per  $\lambda$ dv molecule as compared to 0.5 in previous experiments. When hybridization preparations of these  $\lambda$ dv's with  $\lambda$  DNA were examined, for cases other than  $\lambda$ dv 104, most of the heteroduplexes observed were not of the type expected (Fig. 8 (c)) but, instead, were of the type depicted in Fig. 9. The duplex loops w, are so small (Table 4, 100-300 nucleotides) that sometimes they appear as blibs. Examples are shown in Plates XIII and XIV. This observation indicates that either there is a unique sequence in the  $\lambda$ dv (i. e. they are amphimers instead of inverted dimers) or the heteroduplex is a result of an end effect of an inverted dimer. That is, a few bases at each end of a "snap-back" linear duplex of an inverted dimer cannot be base-paired and probably can nucleate the hybridization with  $\lambda$  DNA. The explanation of the non-inverted sequence proved to be correct for the following reasons.

First, when grids with denatured  $\lambda$ dv's were carefully examined, double-strand, "half-length" (relative to the size of the major open circular species) linears, each with a single-stranded loop at one end, and double-stranded, "full-length" linears, each with either single-stranded loops at both ends or a double-sized, single-stranded loop in the middle, were found. Examples of these molecules are presented in Plate XV. These molecules are expected if a unique sequence does exist at one end of the  $\lambda$ dv's.



(a)  $\lambda_{dv}$  ( except  $\lambda_{dv}$  309 ) /  $\lambda_{c_{26}}$  or parental phage (  $\lambda_{p gal}$  8 vir cII nin or  $\lambda_{gal}$  8 cI 857 nin ) DNA;  $\lambda_{dv309}$  /  $\lambda_{pgal}$  8 vir cII nin .

(c) A branch migrated form of (a).

(d) A branch migrated form of (b).

The notation is the same as in Fig. 1 and Fig. 5 .

Plate XIII Heteroduplexes of type (Fig. ) 5 (d)  $\lambda$ dv and  $\lambda$  DNA's.

A  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 200.

B and D  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 280 .

C  $\lambda$ c<sub>26</sub> /  $\lambda$ dv 280.

Schematic representations are seen in Fig . 9 (a) and (c).

▲ points to the heteroduplex w loop. The small circles are  $\phi$  x 174 and  $\phi$  x 174 RF II.



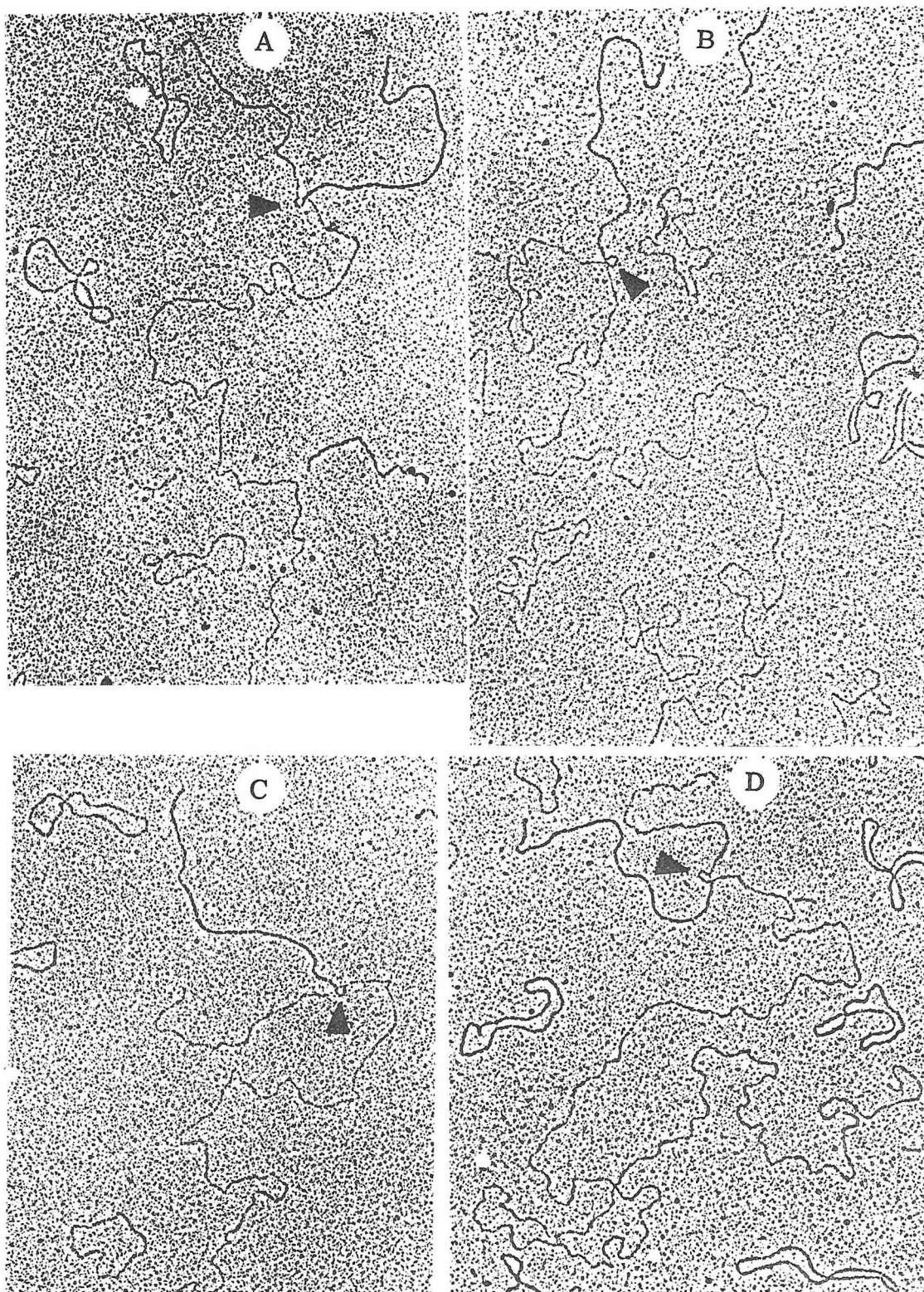


Plate XIII



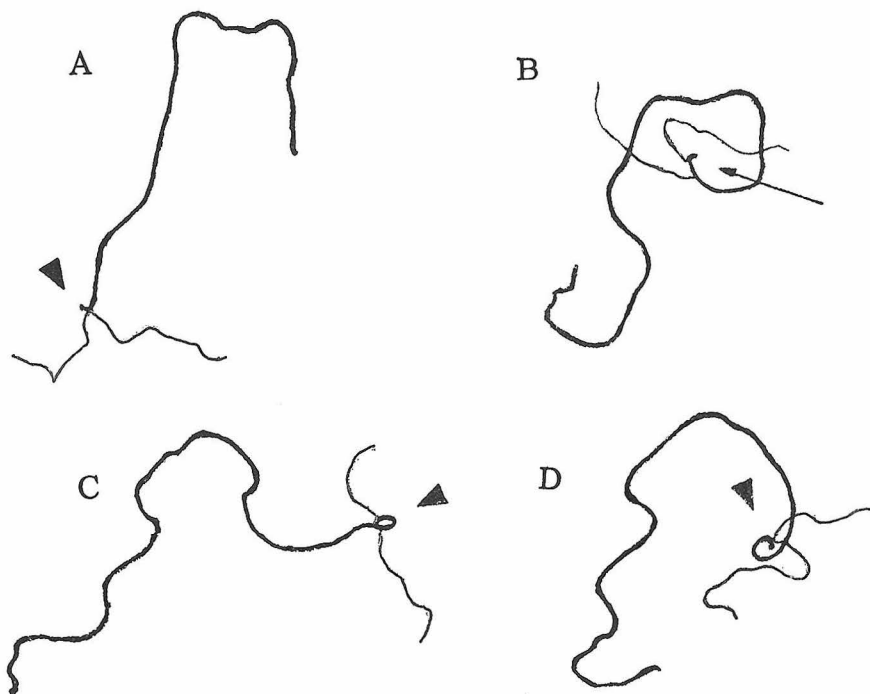


Plate XIV Heteroduplexes of type ( Fig. ) 5 (d)  $\lambda$ dv and  $\lambda$ DNA's.

- A  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 292. Part of the left  $\lambda$  strand is hybridized with a broken  $\lambda$ dv 292.
- B  $\lambda$ c<sub>26</sub> /  $\lambda$ dv 161. Arrow points to the heteroduplex of interest.   
 points to the w loop of a broken snap-back dimer.
- C  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 200.
- D  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 261.

▲ points to the heteroduplex w loop except in B. In A, B, and D, the duplex w loopw are collapsed due to their small sizes. Enlarged illustrative tracings are shown above.

A schematic representation is seen in Fig. 9 (a).

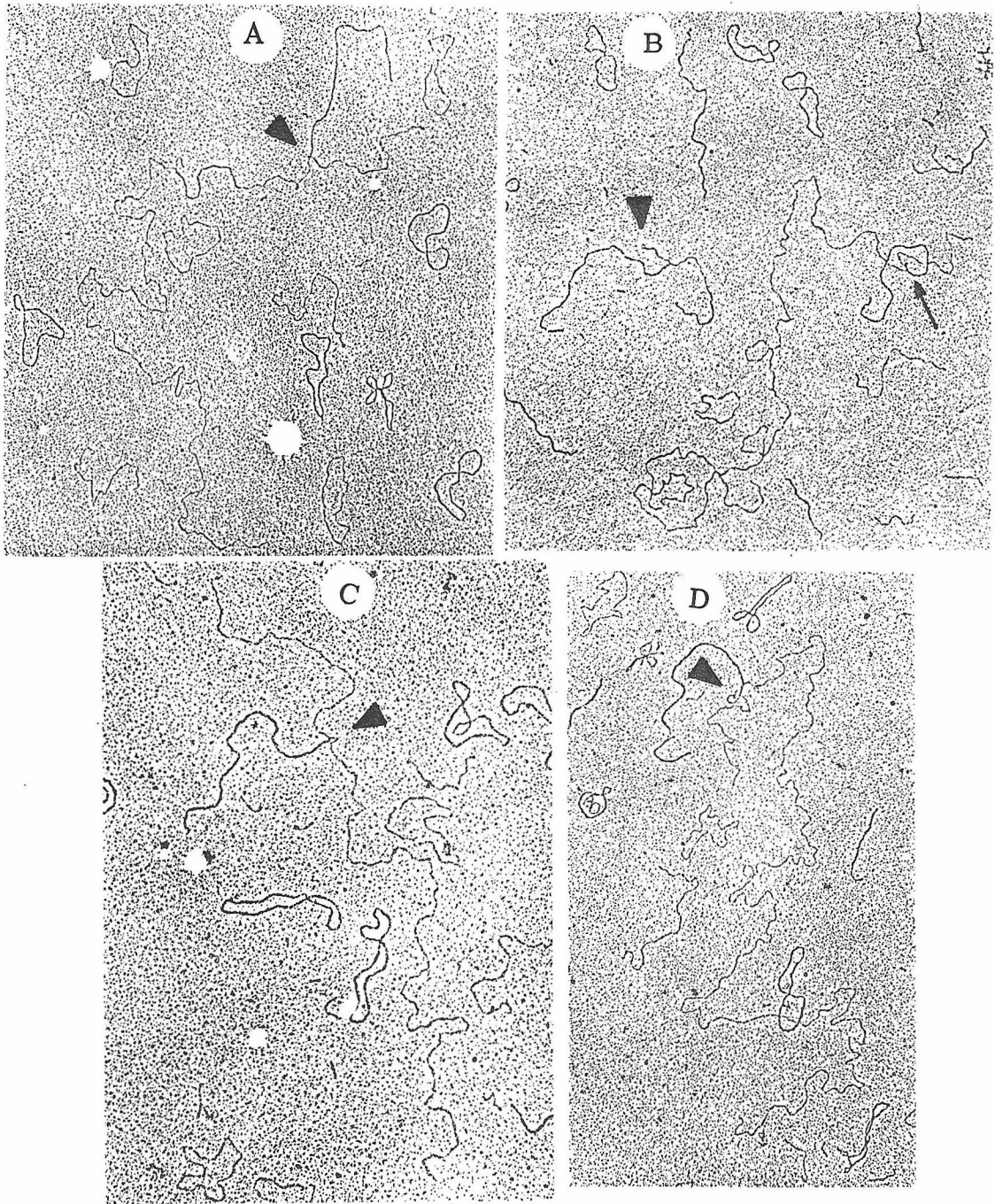


Plate XIV

Plate XV Snap-back  $\lambda$ dv's of type 5(d) showing the w loops

( cf. Fig. 5 )

A, B, and C  $\lambda$ dv 280 diamphimers. A schematic representation is seen in Fig. 5 (dd)' ( the top structure ).

D  $\lambda$ dv 309. A closed circular molecule interlocks a snap-back linear at the w loop .

E  $\lambda$ dv 261 monomer.

F  $\lambda$ dv 280 dimer interlocked with monomer at the w loop.

G  $\lambda$ dv 309 dimer.

H  $\lambda$ dv 309 monomer.

I  $\lambda$ dv 292 monomer.

▲ points to the w loop.

The small circles are  $\phi$  x 174 and  $\phi$  x 174 RF II.

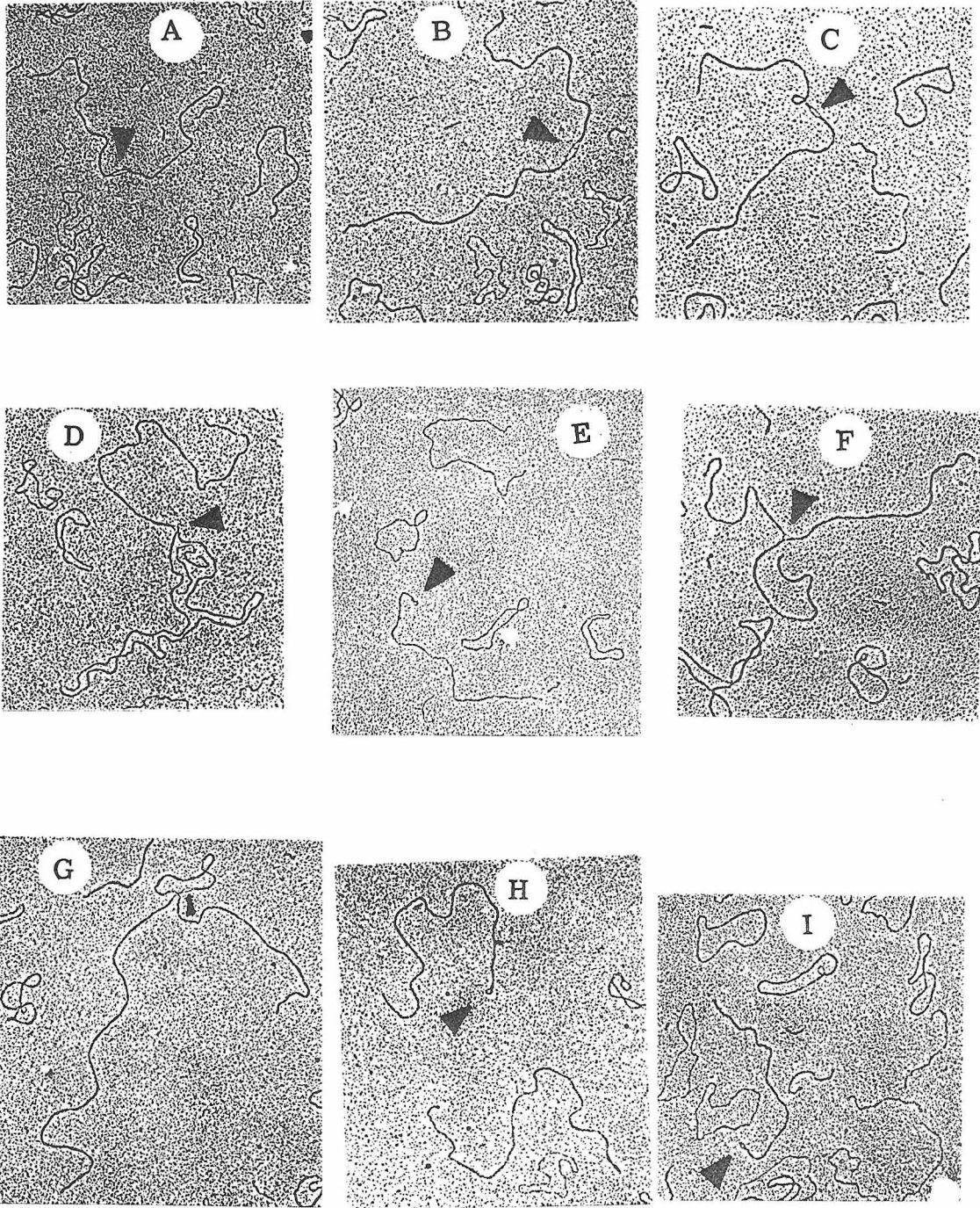


Plate XV

Secondly, careful length measurements indicate: (1) the duplex loop w, (Fig. 9 (a)) has a distinct size for each  $\lambda dv$ , (2) the structure of Fig. 9 (c) is interpreted as a branch-migrated product of the structure of Fig. 9 (a). Duplex zw is not equal in size to the single strand z', as would have been observed if the heteroduplexes had resulted from the "end effect"; instead, (duplex zw - single strand z') always gives a length equal to that of the duplex loop w, regardless of the sizes of zw and z' that vary with the extent of branch migration. Therefore, these  $\lambda dv$ 's are reclassified as amphimers of type 5 (d). The frequency of finding a denatured monomer with the w loop at one end is low; it increases with the loop size. The reason for this is that w is "collapsed" in most of the molecules due to its small size. Since w is located at the end of a linear molecule, the entire molecule appears as a smooth double-stranded linear. Ordinarily, a collapsed loop of 100-300 nucleotides long appears as a blib and is visible if it is situated in the middle of a double-stranded linear molecule. However, I cannot exclude the possibility that the preparation is heterogeneous with most of the molecules having lost the unique sequence, w.

Type 5 (c)  $\lambda dv$  is interesting in that it has the characteristics of both types 5 (b) and (d)  $\lambda dv$ 's. It has two unique sequences of unequal size. The short one on the nin deletion end, the long one on the immunity region end. Heteroduplexes with  $\lambda$  DNA's are presented in Plate X.

The unique sequences of type 5 (d)  $\lambda$ dv's are all located to the right of the nin deletion except that of  $\lambda$ dv 309. The unique sequence of  $\lambda$ dv 309 extends across the nin deletion so that the heteroduplex of this sequence with  $\lambda^+$  DNA shows a single-stranded nin deletion loop in the duplex loop w (Fig. 9 (b)). Heteroduplexes with its parental phage do not show the nin loop, as expected. Examples are presented in Plate XVI.  $\lambda$ dv 309, derived from  $\lambda$  p gal bio 69 nin is the biggest of all  $\lambda$ dv's studied. The physical size of the monomeric amphimer is about 40.3% of  $\lambda$ , and the genetic size is 20.6% of  $\lambda$ . Therefore, it carries some bio sequence on the left end (Table 4).

### 3. $\lambda$ dv 120.

$\lambda$ dv 120 was made according to the procedures diagrammed in Fig. 10. This  $\lambda$ dv is interesting not only because it contains bacterial genes of the galactose (gal) operon, but also because it is a non-inverted partial dimer. The major DNA species isolated by me is a circular molecule containing two sets of the  $\lambda$  gene complement and one set of the gal genes. In self-renatured samples, minor species containing one or two sets each of the  $\lambda$  and the gal genes and minor species containing two sets of the  $\lambda$  genes only are seen. This result is confirmed by restriction enzyme RI treatment. Endonuclease RI has one cutting site in each  $\lambda$  gene complement of the  $\lambda$ dv's and has none in the gal genes.  $\lambda$ dv 120 prepared in this laboratory was treated with RI endonuclease by J. E. Mertz at Stanford University. The products were duplex linear molecules of

Plate XVI Heteroduplexes of  $\lambda$  and  $\lambda$ dv 309 DNA's.

A  $\lambda$ c<sub>26</sub> /  $\lambda$ dv 309.

B  $\lambda$ p gal 8 vir cII nin /  $\lambda$ dv 309 .

Schematic representations are seen in Fig. 9 (d) and (c) .

▲ points to the heteroduplex w loop; Δ points to the nin deletion loop.



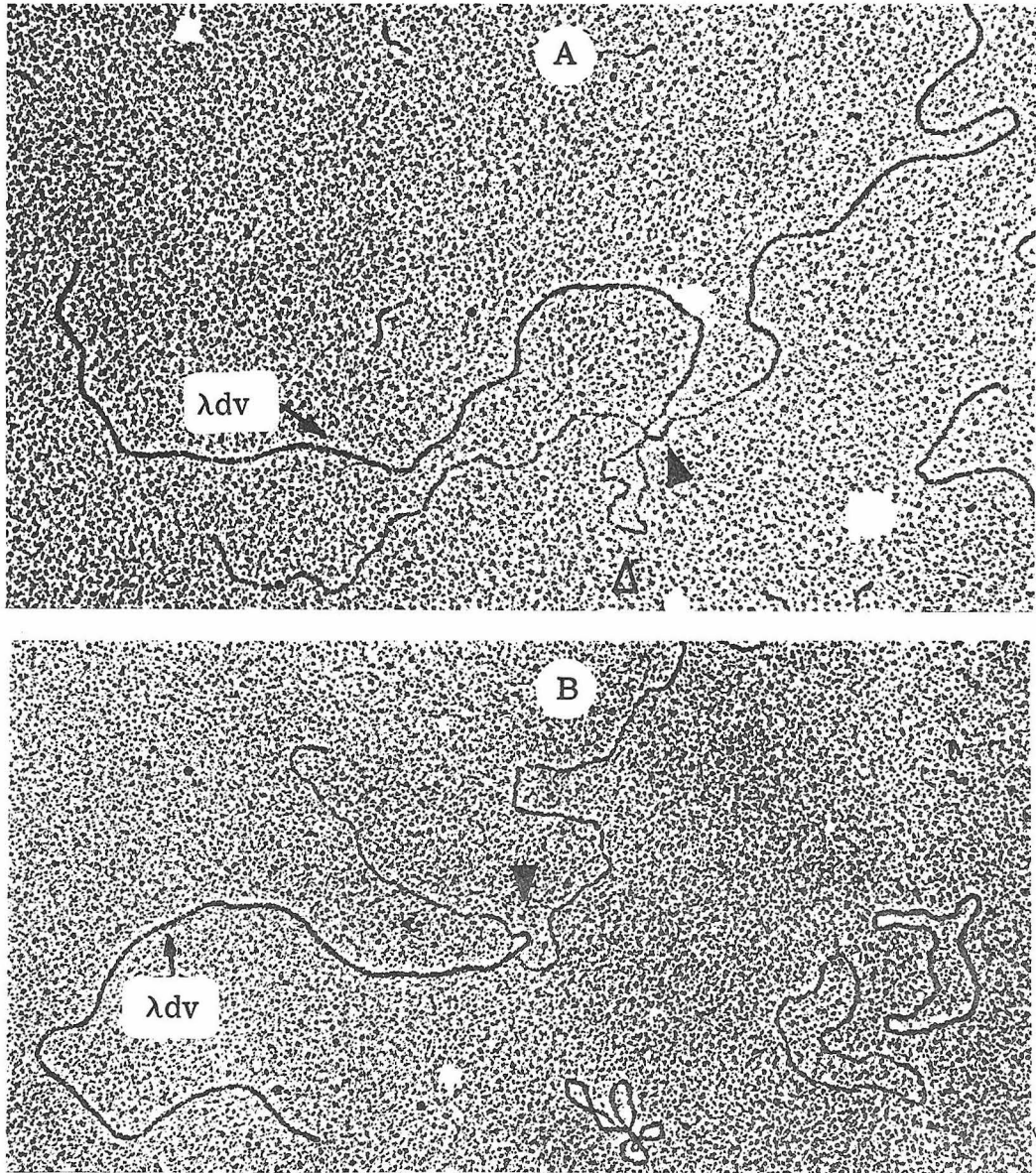


Plate XVI



two sizes corresponding to those of  $\lambda$  genes and  $\lambda$  plus gal genes of the  $\lambda$ dv. A  $\lambda$ dv120 consisting of two sets each of the  $\lambda$  complement and the gal genes has been prepared at Stanford and is cut by RI endonuclease into half-size linear molecules (J. E. Mertz, personal communication). The difference between these two  $\lambda$ dv120 preparations is probably due to the instability of the gal genes. More than 95% of phage  $\lambda$ <sub>b<sub>221</sub> c<sub>I</sub></sub> 857 p g q 4, the parental phage of  $\lambda$ dv120 isolated by Michael Feiss, grown by us were gal<sup>-</sup> segregant. The loss of the gal genes from either the phage or the  $\lambda$ dv is most likely due to the presence of a duplicated DNA sequence on either side of the gal genes. According to M. Feiss, the duplicated segment is less than 0.5% of  $\lambda$ <sup>+</sup> and consists of a small portion of the DNA in the  $\lambda$  P - Q gene interval.

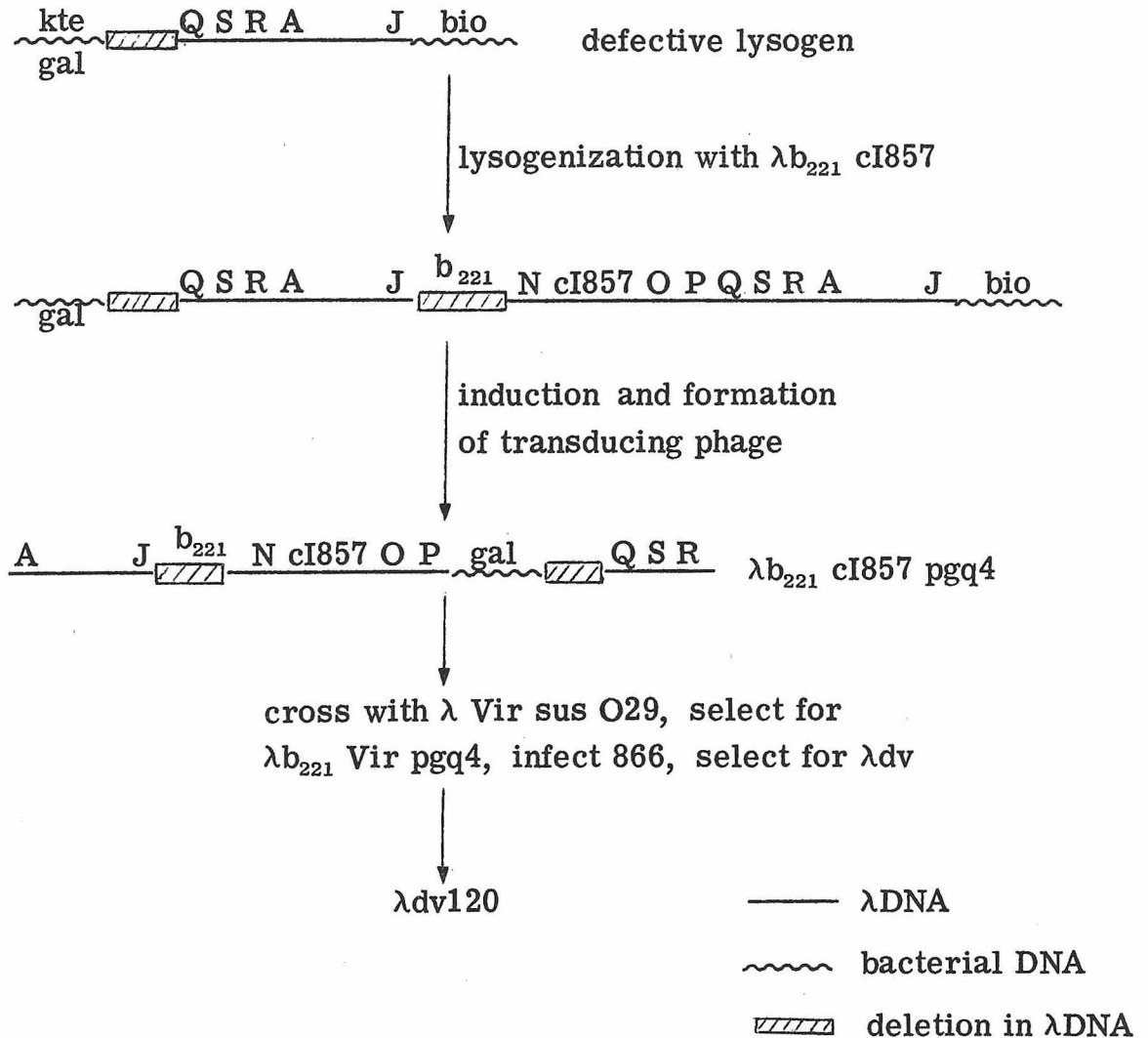


Fig. 10 Formation of  $\lambda$ dv120 ( D.E. Berg, personal communication )

## Further Discussion

### Instability and Heterogeneity of $\lambda$ dv DNAs

$\lambda$ dv is lost spontaneously from its carrier at low frequency. The carrier can also be cured of it by acridine orange treatment. In my hands,  $\lambda$ dv's do not seem to be very stable. Cells that have lost the  $\lambda$ dv were frequently encountered in the single colony test purification procedure. Stabs sometimes completely lose their  $\lambda$ dv upon storage at 4° C. Many other  $\lambda$ dv's were not recovered from cells grown from clones which gave a  $\lambda$ dv positive response in the single colony test. The instability of  $\lambda$ dv is also expressed by the heterogeneity of the  $\lambda$ dv preparations.

Preparations of  $\lambda$ dv's of types 5 (b), (c) and (d) were usually somewhat heterogeneous (Table 4) in length and formed heteroduplexes with  $\lambda$  DNA of different structures at low frequencies. In addition to the main  $\lambda$ dv species, the samples contained: (1) non-inverted, odd-sized molecules, (2) very small (3-4% of  $\lambda^+$ ) circular molecules (It is not known whether they contain  $\lambda$  sequences), (3) molecules that have lost the non-inverted sequence w, in type 5 (c) and probably type 5 (d)  $\lambda$ dv's, (4) molecules that have lost the non-inverted sequences w and a fraction of the neighboring inverted sequences in types 5 (c)  $\lambda$ dv.

Irregular molecules have also been observed in preparations of non-inverted  $\lambda$ dv's derived from nin<sup>-</sup> phage (Table 4). They are  $\lambda$ dv molecules with deletions as indicated by shortened heteroduplexes with  $\lambda$  DNA and deletion loops in self-renatured samples.

Correlation between the sizes and the positions of the  
unique sequences of the amphimers and the frequencies of  
the heteroduplexes formed between such sequences and  
 $\lambda$  DNA

An interesting correlation was observed between the size of the unique sequence  $\underline{w}$ , (Fig. 5) and the frequency of heteroduplexes depicted in Figs. 6 (c), (d) and 9. The smaller the loop  $\underline{w}$ , the lower the frequency of the heteroduplex. An explanation for this observation is that the single-stranded  $\lambda$  DNA has to "thread through" the loop  $\underline{w}$ , in order to pair with it. As  $\underline{w}$  decreases, it becomes more and more difficult for the  $\lambda$  DNA to do so, until, finally,  $\lambda$  DNA fails to thread through. Therefore, the failure of  $\lambda$ dv104 to form heteroduplexes of this type with  $\lambda$  DNA might be either due to a complete absence of the unique sequence or due to an extremely small size of such a sequence.

There is also a correlation between the position of the non-inverted sequence and the frequency of the heteroduplex formed. A heteroduplex with an "outer" loop  $\underline{w}$ , as in Fig. 6 (c) is seen at a higher frequency than that of Fig. 6 (d) with an "inner" loop  $\underline{uv}$ . This is attributed to the excluded volume effect: the ends of a DNA molecule are more exposed than the internal segments and thus, are more accessible for heteroduplex formation.

Comparison of genetic data obtained from recombination  
with physical data obtained from electron microscopy

The electron microscope mapping results as to the sequences present in the  $\lambda$ dv's are in agreement with the genetic results except for the case of  $\lambda$ dv 266 (underlined in Tables 3 and 4). The discrepancy may have arisen from a reduction in size of the  $\lambda$ dv DNA after the genetic tests were performed due to the instability of the  $\lambda$ dv.

Aside from the discovery of the inverted dimer and the amphimers, electron microscopy proved to be a more sensitive method for mapping than genetic tests in two cases. Genetic analyses indicated that  $\lambda$ dv 204 DK does not extend beyond the nin deletion; however, the heteroduplex of the  $\lambda$ dv and  $\lambda^+$  DNA clearly shows the nin deletion loop, and the duplex ends 200 base pairs to the right of the nin loop (Plate IV).  $\lambda$ dv 309 was first thought to be able to fit in a  $\lambda$  phage with a 5% deletion in the compound phage test; yet electron microscopy data show that the  $\lambda$ dv 309 monoamphimer is 40.3% of  $\lambda^+$ . By further genetic tests, it was then observed that the "compound phage" is actually an i <sup>$\lambda$</sup>  recombinant and is not heterozygous for i<sup>21</sup>/i <sup>$\lambda$</sup> , as a compound phage should be.

Hot regions for recombination

Table 4 shows that the right end points of many of the  $\lambda$ dv's derived from  $\lambda$  nin<sup>-</sup> lie in the immediate vicinity of the nin deletion, and that their left end points lie close to the left boundaries of i<sup>434</sup>

and  $i^{\lambda}$ . Our electron microscope mapping data were obtained either by measuring from the right end points (REP's) to the left end points (LEP's) (mapping methods A and B-1) or by using the nin deletion which is very close to the REP's of the  $\lambda$ dv's as a reference point (method B-2). Consequently, the standard deviations for the REP's (0.1 to 0.2% of  $\lambda$ ) are smaller than those for the LEP's (usually smaller than 0.5%) We therefore believe that the right end points of these  $\lambda$ dv's are located very close to one another, yet they are not identical, whereas the left end points that are mapped as being close but different may really be identical. In any case, the excision recombination frequency in these two regions is rather high.

#### Inversions and the nin deletion

No function has yet been attributed to the DNA segment deleted in the nin deletion (abbreviated as the nin segment) except that the nin deletion enables the  $\lambda$  phage to bypass the requirement for a functional N gene product in phage growth (Court and Campbell, 1972; Fiandt, Hradecna, Lozeron and Szybalski, 1971). However, phages carrying the nin deletion are undoubtedly somewhat defective because, even when they are  $N^+$ , they are slow growing and make very small plaques on plating bacteria. This phenomenon indicates that the nin segment carries an important, yet non-vital function. Since all of the  $\lambda$ dv's carrying inverted repeats originate from  $\lambda$  nin<sup>-</sup> phage and none has been found from  $\lambda$  nin<sup>+</sup> phage, a structural role

of the nin deletion or a functional role of the nin segment leading to abnormal replication or recombination is implicated in the generation of these  $\lambda$ dv's. To examine the possibility of a structural anomaly due to the nin deletion, denatured  $\lambda^+$  and nin<sup>-</sup> DNA ( $\lambda$  gal 8 vir cII nin) was mounted under a mild denaturing condition (30% onto 5% formamide). No peculiar structure (such as a hair-pin inversion with or without a single-stranded loop connecting the inverted sequences) was found in either DNA. However, this negative result by no means rules out the existence of an unusual structure that is "melted out" even under this mild mounting condition used. If  $\lambda$ dv's are indeed formed by recombination during  $\lambda$  DNA replication as postulated by Berg and Kaiser (1972), then the functional role of the nin segment could very well be involved in DNA replication. For instance, the nin segment could specify one or more proteins that would help stabilize the replication intermediate and assure efficient and accurate replication.

$\lambda$ dv 249 is worth further commenting in this connection:

The head-to-head inverted dimer of the  $\lambda$ dv 249 monoamphimer, which already contains an inverted sequence, could be formed by a recombinational event between a plus and a minus strand of two monoamphimers in the inverted, duplicated region or by an abnormal replication process of the monoamphimer. The existence of such inverted dimers, together with tandem dimers and trimers containing tandem and inverted sequences in the same preparation, suggests inversion can be produced by erroneous recombination or replication

after the production of the  $\lambda$ dv. The hypotheses concerning nin are consistent with either process that produces the mixed population of  $\lambda$ dv 249.

Whatever the mechanism that produces inverted  $\lambda$ dv's, three facts have to be accounted for: (1) all  $\lambda$ dv's containing inverted repeats except for  $\lambda$ dv's 104 and 249 have non-inverted, non-duplicated right ends located very close to the nin deletion, (2) the left boundaries of the inverted sequences vary widely among the different  $\lambda$ dv's, (3) there are non-inverted, non-duplicated sequences beyond the left boundaries of the inverted sequences in some of the  $\lambda$ dv's ( $\lambda$ dv's 249 and 204 LC).

In conclusion, whatever the role of the nin deletion or the nin segment, I believe that the generation of  $\lambda$ dv's containing inverted repeats is connected in some way to the nin deletion by recombination or replication during or after the formation of the  $\lambda$ dv's.



## References:

- Bellett, A.J.D., Busse, H.G., and Baldwin R.L. (1971) In The Bacteriophage Lambda, ed. by A.D.Hershey, p.501  
New York: Cold Spring Harbor Laboratory.
- Berg, D.E. (1971) In The Bacteriophage Lambda, ed.by A.D.Hershey,  
Hershey, p.667. New York: Cold Spring Harbor Laboratory.
- Berg, D.E., and Kaiser, A.D.(1972) In press.
- Busse, H.G., and Bladwin, R.L.(1972) J.Mol.Biol., 65, 401.
- Champoux, J.J., (1969) Ph.D.Dissertation, Stanford University.
- Chow, L.T., Boice, L., and Davidson, N.(1972) J.Mol.Bio. 68,  
391.
- Chow, L.T., and Davidson, N.(1973) J.Mol.Biol., in press.
- Court, D., and Campbell, A. (1972) J.Virol. in press.
- Davidson, N. and Szybalski, W.(1971) In The Bacteriophage Lambda,  
ed.by A.D.Hershey, p.45. New York: Cold Spring Harbor  
Laboratory.
- Davis, R.W., Simon, M.N., and Dacidson, N. (1971) In Methods in  
Enzymology, vol. 21D, p.413.
- Fiantt, M., Hradecna, Z., Lozeron, H., and Szybalski, W. (1971)  
In The Bacteriophage Lambda, ed.by A.D.Hershey, p.329.  
New York: Cold Spring Harbor Laboratory.
- Franklin, N.C., (1971) In The Bacteriophage Lambda, ed. by A.D.  
Hershey, p.175. New York: Cold Spring Harbor Laboratory.
- Hobom, G., and Hogness, D.S. (1972) J.Mol.Biol., inpress.
- Inman, R.B. (1966) J.Mol. Biol., 18, 464.
- Matsubara, K., and Kaiser, A.D. (1968) Cold Spring Harbor Symp.,

Quant. Biol. 33, 769.

Schös, M., and Inman, R. B. (1970) J. Mo. Biol., 51, 61

Sharp, P. A., Hsu, M-T., Ohtsubo, E., and Davidson, N., (1972)  
Proc. US Nat. Acad. Sci., in press.

Simon, M. N., Davis, R. W., and Davidson, N., (1971) In The  
Bacteriophage Lambda, ed. by A. D. Hershey, p. 313.  
New York: Cold Spring Harbor Laboratory .

Spiegelman, W. G. (1971) Virol. 43, 16.

Stevens, W. F., Adhya, S., and Szybalski, W. (1971) In The  
Bacteriophage Lambda, ed. by A. D. Hershey, p. 515.  
New York: Cold Spring Harbor Laboratory.

## PROPOSITION I

MAPPING OF SV40 mRNA SYNTHESIZED IN VIVO

SV40 RNA's transcribed from both SV40 DNA strands can be extracted from infected cells. This proposition is to map these RNA's, making use of the SV40 DNA fragments generated by restriction endonucleases that makes specific cuts in the SV40 DNA. Terminal deoxyribonucleotidyl transferase will be used to aid the mapping.

SV40 specific RNA can be isolated from infected permissive ( lytic response ) hosts and from non-permissive ( transformed ) hosts (1-3). Based on nitrocellulose filter hybridization, 1/3 of the SV40 genome is transcribed from one strand of the DNA shortly after lytic infection before DNA replication. These RNA species are called early RNA. Some of the early RNA is also found in transformed cells. Late in lytic infection, another class of RNA complementary to 2/3 of the genome is made, in addition to the continued production of the early RNA (2, 4). Some of the RNA extracted late in infection are complementary to one another (5). It is not known where on each of the DNA strand the RNA's are transcribed.

Direct mapping of SV40 RNA's by DNA-RNA hybridization and electron microscopy has several drawbacks : (1) it is difficult to isolate the needed large quantities of SV40 RNA from infected cells. (2) because SV40 DNA is a double stranded circular molecule, it

has no immediate point of reference for electron microscope mapping. To provide a unique reference point against which early and late RNA species can be mapped, I propose to use restriction endonucleases to introduce defined breaks into the circular genome .

Several bacterial strains contain restriction endonucleases which make double stranded scissions at specific nucleotide sequences in DNA not properly methylated (6-13). SV40 DNA is cut into a set of 11 unique fragments of different lengths (molecular weight,  $0.74-7.2 \times 10^5$  daltons ) by endonuclease R from *Haemophilus influenzae* (13), whereas only one unique double stranded break is introduced by endonuclease R I ( coded by the drug resistant factor R I ) to give a linear molecule ( 14, 15 ).

The experiments proposed are:

( 1 ) Enzymatic scission of SV40 DNA by restriction endonucleases R and RI.

The "R" fragments will be separated by acrylamide gel electrophoresis.

( 2 ) Mapping of the "R" fragments on the "RI" linear molecule by electron microscope heteroduplex method.

Fragment 1 produced by endonuclease R can be positioned on the whole linear SV40 prepared by endonuclease RI cleavage. Mapping of the other DNA fragments should be carried out by co-hybridization with fragment 1 ( or another previously mapped fragment ) as an internal reference point.

At the lower limit, double stranded heteroduplexes of 200-400 bases without single stranded fork at one or both ends can be

distinguished from single stranded DNA. Some of the fragments produced by endonuclease R are too short to be visible. Terminal deoxynucleotidyl transferase may be used to extend such short pieces of DNA (16). This enzyme polymerizes deoxyribonucleoside triphosphates into polymers using a 3'-OH terminus of a DNA strand as a primer (Equation 1). It polymerizes dATP most efficiently,



with  $n$  as long as 400-500 residues. SV40 fragments lengthened by the terminal transferase can be hybridized to the linear wholes (Fig. 1 a and b). The single stranded tails of poly dA indicate the 3' end of the fragment (population 1a will not be distinguishable from population 1b when the fragment itself is short). The Tails can be further expanded and made more visible by cohybridizing with poly (dT) $_{n>500}$  (Fig. 1c)

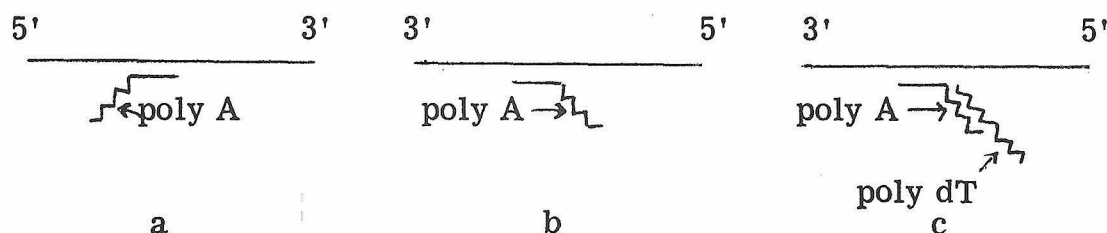


Fig 1

In the general sense, this system can be used to map any small pieces of DNA on a completely homologous DNA.

(3) "Mapping" of SV40 RNA molecules indirectly by DNA-RNA membrane filter hybridization.

The isolation of SV40 specific RNA from transformed or lytically infected cells is facilitated by the recent finding that all SV40 specific RNA molecules have polyA added to their 3' ends(17).

They can be obtained in highly pure form by sequential hybridization to poly U and denatured SV40 DNA immobilized on filter. The SV40 mRNA's thus obtained will be tested against the various "R" fragments by membrane filter hybridization. Since the DNA fragments will have already been positioned on the SV40 linear wholes, the RNA molecules are correspondingly mapped on the linear SV40 DNA.

( 4 ) Identification of the genetic message carried by the fragments.

It may be possible to test for genes in some of the large fragments. Susceptible cells in tissue culture can be infected by linear (15), and circular monomeric (18), dimeric or polymeric SV40 DNA's (19). I propose to test whether such cells can take up linear or cyclized SV40 fragments. Cyclization can be accomplished by (a) adding polydA and polydT to different populations of the same fragments, using the terminal transferase. (b) denaturation and re-naturation of the 2 populations. Cyclization will occur when a fragment with a polydA tail hybridizes with another fragment with a polyd T tail. (c) filling and sealing the gaps ( generated by the uneven length of the added polydA and polydT ) by the large subunit of the E. coli DNA polymerase (20) and DNA ligase (21, 22). The cells "infected" by the fragment singly or in various combinations will then be examined for the expression of the various viral functions, for instance, the production of antigens, loss of contact inhibitions, and lysis of cells. It would also be interesting to test whether one of the fragments can exist autonomously in animal cell as  $\lambda$ dv and other plasmids can in bacterial cells. Such a fragment presumable contains the origin of replication (23-25).

## References:

- 1 Khoury, G., and Martin, M.A. (1972) *Nature New Biol.* 238, 4
- 2 Khoury, G., Byrne, J.C., and Martin, M.A. (1972) *Proc. US Nat. Acad. Sci.*, 69, 1925
- 3 Lindberg, U., and Darnell, J.E. (1970) *Proc. US Nat. Acad. Sci.*, 65, 1089
- 4 Sambrook, J., Sharp, P.A., and Keller, W. (1972) *J. Mol. Biol.* 70, 57
- 5 Aloni, J. (1972) *Proc. US Nat. Acad. Sci.*, 69, 2404
- 6 Kakano, T., and Watanabe, T. (1966) *Biochem. Biophys. Res. Comm.*, 25, 192
- 7 Boyer, H., Scibienski, E., Slocum, H., and Roulland-Dussoix, D., (1971) *Virology*, 46, 703
- 8 Linn, S. and Arber, W. (1968) *Proc. US Nat. Acad. Sci.*, 59, 1300
- 9 Meselson, M., and Yuan, R. (1968) *Nature*, 217, 1110
- 10 Roulland-Dussoix, D., and Boyer, H.W. (1969) *Biochim. Biophys. Acta.*, 195, 219
- 11 Amith, H.O., and Wilcox, K.W., (1970) *J. Mol. Biol.* 51, 379
- 12 Adler, S.P., and Nathans, D. (1970) *Fed. Proc.*, 29, 725
- 13 Danna, K., and Nathans, D. (1971) *Proc, US Nat. Acad. Sci.*, 68, 2913
- 14 Morrow, J. F. and Berg, P. (1972) In press.
- 15 Mertz, J.E., and Davis, R.W. (1972) In press.
- 16 Kato, K., Gonclaves, J.M., Houts, G.E., and Bollum, F.J., (1967) *J. Biol. Chem.*, 242, 2780
- 17 Weinberg, R.A., Ben-Ishai, Z., and Newbold, J.E., (1972)

- Nature New Biol., 238, 111.
- 18 Jaenisch, R., and Levine, A.J. (1971) J. Mol. Biol., 61, 735.
  - 19 Dubbs, D.R., Kit, S., Jaenisch, R., and Levine, A.J. (1972)  
J. Virol., 9, 717.
  - 20 Brutlag, D., and Kornberg, A. (1972) J. Biol. Chem., 247, 224.
  - 21 Olivera, B.M., and Lehman, I.R. (1967) Proc. US Nat. Acad. Sci.,  
57, 1426.
  - 22 Olivera, B.M., and Lehman, I.R. (1967) Proc. US Nat. Acad. Sci.,  
57, 1700.
  - 23 Nathans, D., and Danna, K.J. (1972) Nature New Biol., 236, 220.
  - 24 Thoren, M.M., Sebring, E.D., and Salzman, N.P. (1972)  
J. Virol., 10, 462.
  - 25 Fareed, G.C., Garon, C.F., and Salzman, N.P. (1972)  
J. Virol., 10, 484.

Added note: After this proposition is written , Jackson, Symons, and Berg ( Proc. Nat. Acad. Sci., 69, 2904 , (1972) ) reported that they have used the termianl transferase, DNA polymerase and DNA ligase to creat SV40 dimers and SV40 containing  $\lambda$ dv 120. Their success lends support to the mapping experiments in this proposition.



## PROPOSITION II

Gene Homology Between Bacteriophages S13 And  $\phi$ X174

$\phi$ X174 RF DNA is fragmented by restriction endonucleases R and Z into discrete pieces. Genetic markers have been rescued from some of the fragments. The sizes of the  $\phi$ X174 cistrons have been determined and the direction of translation is found to be the same as the genetic map. Phage S13 is known to be closely related to  $\phi$ X174 in terms of morphology, serological interaction, genetic recombination and complementation. Experiments are proposed to use the above mentioned information, the electron microscope heteroduplex method and  $T_m$  depression to study their gene homology.

The close relatedness of bacteriophages S13 and  $\phi$ X174 is evidenced by their morphology, serological interaction ( Zahler, 1958 ), genetic recombination and complementation ( Tessman, and Shleser, 1963 ). Genomes of both phages are single stranded circular DNA molecules of  $1.7-1.8 \times 10^6$  daltons molecular weight ( Spencer, Cerny, Cerna, and Delaney, 1972 ). Their coat proteins and certain other phage specific proteins found in infected cells have very similar mobilities on acrylamide gel electrophoresis ( Poljak, and Suruda, 1969; Jeng, Gelfand, Hayashi, Shleser, and Tessman 1970 ). Amino acid analyses of their coat proteins are almost identical ( Poljak et al., 1969 ). However, S13 DNA is slightly richer in the proportion of pyrimidine nucleotides ( Spencer et al., 1972 ). For these reasons,

the heterduplexes formed between the two phage DNA's may show regions of complete homology and of complete inhomology as in the case of lambdoid phages ( Simon, Davis, and Davidson, 1971; Fiandt, Hradecna, Lozeron, and Szybalski, 1971 ) or partial inhomology as in the case of  $T_3$  and  $T_7$  ( Davis, and Hyman, 1971 ) when examined by the electron microscope.

A recent determination of the sizes and relative order of the  $\phi X174$  cistrons ( Benbow, Mayol, Picchi, and Sinsheimer, 1972 ) and the availability of *Haemophilus influenzae* restriction enzyme ( endonuclease R ) ( Edgell, Hutchison, and Sclair, 1972 ) and *Haemophilus aegyptius* restriction enzyme ( endonuclease Z ) ( Middleton, Edgell, and Hutchison, 1972 ) that generate discrete fragments from  $\phi X174$  RF DNA now make it possible to compare directly the gene homology of the two phages.

Endonuclease R and Z have different site specificities; the former cuts  $\phi X174$  RF DNA into a set of at least 13 fragments, whereas the latter produces a different set of at least 11 fragments. Both sets have been separated and sized by Edgell, Hutchison and associates ( Tables 1 and 2 of appendix ). Genetic markers have been rescued from some of the fragments; for instance, the wild type allele for the mutation  $F_{ts}41D$  ( cistron F ) is in fragments  $Z_1$  and  $R_1$ , and that for mutation  $B_{ts}9$  ( cistron B ) is in fragments  $Z_3$  and  $R_5$ .

The direction of the translation of the  $\phi X174$  genome is D-E-(J)-F-G-H-A-B-C, and is identical to the order of the genetic map which is circular with C and D joined ( Benbow, Hutchison, Fabricant, and Sinsheimer, 1971 ). The sizes of the cistron have

been estimated from the molecular weights of the proteins encoded by each of them ( Table 3 of Appendix )

From data summarized in Tables 1, 2, and 3, a direct comparison of gene homology of the phages can be studied by the following set of experiments:

First we will map the fragments of  $\phi$ X174 RF produced by endonuclease Z on  $\phi$ X174 viral DNA as follows:

- (a) Separation and isolation the "Z" fragments.
- (b) Mapping of fragments  $Z_1$  and  $Z_3$  on viral DNA by the electron microscope heteroduplex method. From Tables 1, 2, and 3,  $Z_1$  and  $Z_3$  are expected to be about  $180^\circ$  from each other; therefore it is not possible to orient them relative to the genetic map. To accomplish this, another large fragment,  $Z_2$ , should be tested for its ability to rescue markers in other cistrons. Fragment  $Z_2$  is then physically mapped on the viral DNA relative to  $Z_1$  and  $Z_3$ . Fragments  $Z_1$  and  $Z_3$  are larger than the cistrons whose markers they can rescue. However, markers can be rescued from fragments which do not necessarily contain the whole cistron. To refine the map, fragments  $R_1$  and  $R_5$  which can rescue the same two markers as  $Z_1$  and  $Z_3$  should be mapped by constructing the following heteroduplexes :  $R_1+Z_3+\phi$ Xss ( ss stands for single strand );  $R_1+R_5+\phi$ Xss or  $Z_1+R_5+\phi$ Xss. The regions covered by  $R_1$  and  $Z_1$  and by  $R_5$  and  $Z_3$  give the more accurate positions of markers  $Ft_s 41D$  and  $Bt_s 9$  ( Fig. 1 )

Secondly we will study the structure of the heteroduplex of "whole" S13 and  $\phi$ X174 DNA's. Because of the close relatedness of the two phages, heteroduplex made from  $\phi$ X174 RFII and S13 RF II

or S13 viral DNA ( and vice versa ) may show regions of homology and inhomology with scattered substitution and/or deletion loops as depicted in Fig. 2. If this is true, then structure such as that

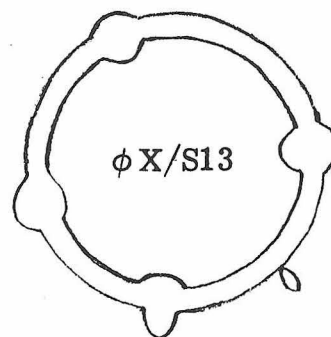
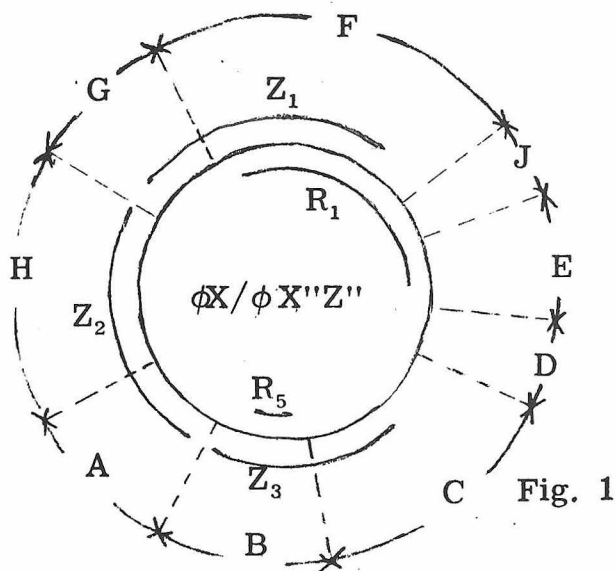


Fig. 2

of Fig. 3 will be observed when heteroduplexed of S13 viral DNA and  $\phi X174 "Z"$  fragments are prepared.

S13/ $\phi X "Z"$

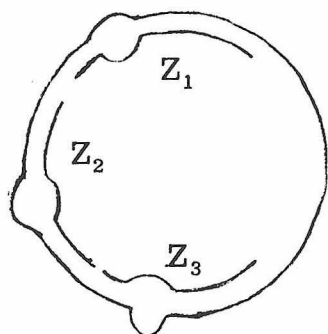


Fig. 3

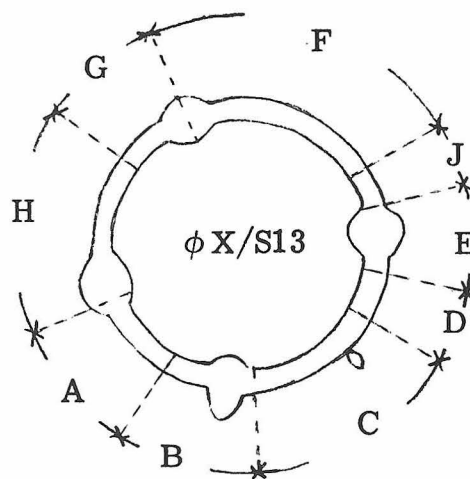


Fig. 4

From these experiments; it is possible to deduce the gene homology between S13 and  $\phi X174$ . The gene homology map of the two phages showing different genes with different extent of homology

is obtained by superimposing Figures 1, 2, and 3 ( Fig. 4).

Comparison of the gene homology map with recombination data and protein amino acid analyses can then be made. Because the coat proteins of the two phages ( coded by genes F, G, H, J, and B in  $\phi$ X174 ) are almost identical, nucleotide inhomology would indicate that different nucleotide sequences could code for almost identical proteins, as permitted by the code degeneracy. One can also find out whether functional proteins can be produced by recombination in which somewhat different DNA segments have been exchanged.

It may be, however, that the homology of  $\phi$ X174 and S13 is so high that the heteroduplex depicted in Fig. 2 is not observed even under highly denatured condition, but appears as double stranded throughout as in the case of *M. luteus* phages  $N_1$  and  $N_6$  ( Lee, and Davidson, 1970 ). Then the following more sensitive melting point depression experiments should be carried out:

- (1) The same as experiments (a) and (b) discussed above except that all the "Z" fragments will be mapped on the viral  $\phi$ X174 DNA. By combining these physical mapping results and Fig. 1, it is possible to determine the gene(s) contained in each of the fragments. To map some of the small fragments, terminal transferase which has been discussed in Proposition I can be used.
- (2) The melting point of each of the "Z" fragments of  $\phi$ X174 RF is determined.
- (3) The melting point of each of the "heteroduplex fragments" is determined. "The heteroduplex fragment" is prepared by hybridizing S13 RF II or viral DNA with the "Z" fragment of  $\phi$ X174. The unpaired, single

strands at both sides of the heteroduplex is then removed by single strand specific endonuclease ( Carlson, Schmid and Martin, 1972 )

From the depression of melting of each of the heteroduplex fragments as compared to the melting point of the homoduplex fragment, the extent of inhomology of each of the fragments, hence the genes contained, of the two DNA's can then be determined, using the formula that 1.5 % random base mismatches depress the melting point by 1° C ( Laird, McConaughy and McCarthy, 1969 ).

#### Appendix

Table 1  $\phi$ X174 RF fragments produced by restriction endonuclease R

Electrophoresis band	No. of fragments	Size in base pair	marker rescued
R <sub>1</sub>	1	1, 470	Ft <sub>S</sub> 41D
R <sub>2</sub>	1	1, 000	
R <sub>3</sub>	1	760	
R <sub>4</sub>	1	510	
R <sub>5</sub>	1	340	Bt <sub>S</sub> 9
R <sub>6</sub>	3	270	
R <sub>7</sub>	2	220; 200	
R <sub>8</sub>	1	110	
R <sub>9</sub>	1	80	

( From Edgell et al. , 1972 )

Table 2  $\phi$  X174 RF fragments produced by restriction endonuclease Z

Electrophoresis band	No. of fragments	Size in base pair	Marker rescued
Z <sub>1</sub>	1	1690	Ft <sub>S</sub> 41D
Z <sub>2</sub>	1	1350	
Z <sub>3</sub>	1	1025	Bt <sub>S</sub> 9
Z <sub>4</sub>	1	600	
Z <sub>5</sub>	1	215	
Z <sub>6</sub>	2	175	
Z <sub>7</sub>	1	120	
Z <sub>8</sub>	1	90	
Z <sub>9</sub>	1	40	
Z <sub>10</sub>	1	20	

( From Middlton et al. 1972 )Table 3  $\phi$  X174 cistrons ( Benbow et al., 1972 )

Cistron	Measured molecular weight of protein coded ( daltons)	Sizes of * cistron in base pair
D	14, 500	350
E	17, 500	422
F	50, 000	1206
G	20, 500	495
H	37, 000	893
A	14, 000 Or 67, 000	338 or 1616
B	25, 000	603
C	34, 000 or 7, 000	820 or 169
J	9, 000	217
(I)	7, 000 or 34, 000	169 or 820

\* The genome of  $\phi$ X174 is 5,500 bases; the total molecular weight of proteins coded is 228,000; the size of each cistron is obtained by dividing the molecular weight of the protein coded by 41.45.  
 $= 228,000/5,500$ .

#### References:

- Benbow, R. M.; Hutchison, C. A. III; Favricant, J. D., and Sinsheimer, R. L., J. Virol., 7, 549 (1971)
- Benbow, R. M.; Mayol, R. F.; Picchi, J. C.; and Sinsheimer, R. L., J. Virol., 10, 99 (1972)
- Carlson, J.; Schmid, C.; and Martin, K.,  
 Davis, R. W.; and Hyman, R. W., J. Mol. Biol., 62, 287 (1972)
- Edgell, M. H.; Hutchison, C. A. III, Sclair, M. J. Virol., 9, 574 (1972)
- Fian dt, N.; Hradecna, A.; Lozeron, H.; and Szybalski, W. The Bacteriophage  $\lambda$  ed. by Hershey, A. D., p. 329 New York, Cold Spring Harbor Laboratory (1971)
- Jeng, Y.; Gelfand, D.; Hayashi, M.; Shleser, R., and Tessman, E. S. J. Mol. Biol., 49, 521, (1970)
- Laird, C. D.; McConaughy, B. L.; and McCarthy, B. J., Nat, 244, 149 (1969)
- Lee, C. S., and Davidson, N., Virol. 40, 102 (1970)
- Middleton, J. H.; Edgell, M. H., and Hutchison, C. A. III, J. Virol. 10, 42 (1972)
- Poljak, R. O.; and Suruda, A. J., Virol. 39, 145 (1969)
- Simon, M., Davis, R. W., and Davidson, N., The Bacteriophages  $\lambda$  ed. by Hershey, A. D., p. 313 (1971)
- Spencer, J. H.; Cerny, R.; Cerna, E., and Delaney, A. D., J. Virol.



10, 134 ( 1972 )

Tessman, E. S., and Shleser, R., Virology 19, 239 ( 1963 )

Zahler, S. A., J. Bact. 75, 310 ( 1958 )

Added note: Since this proposition is written, Hutchison, Middleton, and Edgell ( Abstract of papers presented at the Bacteriophage Meeting 1972, Cold Spring Harbor Laboratory, New York ) report that they have successfully salvaged wild type markers from most of the R and Z fragments of  $\phi$ X174 RF.

## PROPOSITION III

Transcription And Maturation Of E. coli Transfer RNA And  
Coliphage T<sub>4</sub> Transfer RNA

It is not known how the clustered genes coding for tRNA in E. coli and coliphage T<sub>4</sub> are transcribed and processed. Precursors of individual tRNA which are larger than mature tRNA and are not modified in their bases have been identified. Enzymes involved in the maturation processes of tRNA's from their precursors have been studied in vitro. This proposition is to study the transcription and maturation of the clustered tRNA genes in T<sub>4</sub> and E. coli using T<sub>4</sub> phage s and transducing phages  $\phi$  80 and  $\lambda$  ( as model systems for E. coli) in in vitro system.

Genes coding for tRNA's in Bacillus subtilis ( Smith, Dabnau, Morell, and Marmur 1968; Bleyman, Kondo, Hecht, and Woese 1969 ), E. coli (Cutler, and Evans 1967 ) and coliphage T<sub>4</sub> ( Wilson, Kim, and Abelson 1972 ) are clustered. T<sub>4</sub> has 6-10 tRNA genes located between the lysozyme gene and gene 57 ( Wilson et al. 1972 ). At least two different T<sub>4</sub> tRNA's are transcribed into one single piece of RNA that gives a band in acrylamide gel electrophoresis. A still more slowly moving band suggests the existence of polycistronic tRNA. Transcriptional mapping in Bacillus subtilis indicates that tRNA genes are transcribed polycistronically ( Bleyman et al., 1969 ). In E. coli, there is evidence

for invivo tyrosine tRNA precursors which are longer than the mature tRNA by at least 35-45 nucleotide residues, are not modified in their bases, and have different 3' and 5' ends from the mature tRNA ( Altman, 1971 ). However, there is no evidence for or against polycistronic transcription of tRNA's in E. coli .

Maturation of tRNA is a post-transcriptional event ( Pace, Peterson, and Pace 1970 ), and is defined as the cleavage of the transcription product into the proper length ( Smillie, and Burdon 1970 ), modification of bases ( Fleissner, and Borek 1962, 1963; Weiss, and Legault-Demare 1965; Hayward, and Weiss 1966; Mowshowitz 1970 ), and addition of the 3'terminal -CCAOH sequence ( Preiss, Dieckman, and Berg, 1961; Furth, Hurwitz, Krug, and Alexander 1961; Daniel, and Littauer 1963 ). The latter two classes of processes have been well studied in vitro with partially purified enzymes or cell free extracts. The order of these steps within the maturation process is not known. In the case of HeLa tRNA precursor, cleavage occurs prior to the completion of base modification (Bernhardt, and Darnell 1969; Mowshowitz 1970 ).

This proposition concerns the mode of transcription and maturation of E. coli and  $T_4$  tRNA's .

From in vitro transcription of  $\phi 80\text{psu}_3^+$  ( a transducing phage containing one tyrosine tRNA gene ) DNA by E. coli RNA polymerase, heterogeneous ( 4s to 16s ) tyr tRNA containing species are found on sucrose gradients ( Daniel, Sarid, Beckman, and Littauer 1970 ). Using the same system Ikeda has identified 3 tyr tRNA precursors by acrylamide gel electrophoresis ( Ikeda

1972 ). The precursors have molecular weights of  $2.8 \times 10^5$  (I),  $1.7 \times 10^5$  (II), and  $8.0 \times 10^5$  (III) daltons as determined by acrylamide gel electrophoresis. A crude cell extract from E. coli converts both precursors II and III into RNA species of molecular weight of  $2.8 \times 10^4$  daltons at least 80-90 % of which is of tyr tRNA. The same extract does not hydrolyse mature tRNA. Modifications are probably not prerequisite for such conversion. Although this is a good system to study the maturation of a single tRNA species, it does not provide information as to how a cluster of tRNA genes is transcribed and processed. Phage  $T_4$  with 6-10 tRNA genes clustered together, phage  $\phi 80\text{hpsu}_3^+$  containing two tyrosine tRNA genes ( Russell, Abelson, Landy, Bester, Brenner, and Smith 1970 ) in tandem and phage  $\lambda_{\text{C1857 S}_{\text{t68}}^{\text{h80d gly Tsu}_{36}^{++}}$  containing three tRNA genes in series ( Squires, Conrad Kirschbaum and Carbon 197 ) provide good systems to study these problems in vitro . The  $T_4$  tRNA genes have been mapped by Wilson, Kim and Abelson (1972), the other two phages have been characterized and tRNA genes mapped too. ( Wu and Davidson, 1973; Wu, Davidson, and Carbon, 1972)

Using Ikeda's system, I propose:

- A. to study the transcription of tRNA genes on the three phages in vitro using E. coli RNA polymerase:
  - (a) To separate and size the RNA's made using acrylamide gel electrophoresis.
  - (b) to identify the tRNA's or tRNA precursors by hybridization and competition experiments. The RNA bands from step (a) will be eluted from the gel. Transfer RNA containing species transcribed

from  $\phi 80\text{hpsu}_3^+$  and  $\lambda \text{c}_I857\text{S}_{t68}\text{h80d gly Tsu}_{36}^{++}$  DNA's can be identified as those (1) hybridizable to the template DNA, but not to wild type  $\phi 80$  and wild type  $\lambda$  phage DNA's, and (2) competitive by mature E. coli tRNA in hybridizing to template DNA's. Transfer RNA containing species made on wild type  $T_4$  DNA are identified by comparing gel patterns with those obtained from RNA transcribed on  $T_4$  DNA with tRNA genes deleted ( Wilson et al. 1972 ); these RNA species are competed by  $T_4$  tRNA's in hybridization experiment ( Daniel, Sarid, and Littauer 1970 ).

B. To study the maturation of tRNA:

If tRNA precursors are indeed found, their maturation can then be investigated. From kinetic studies, possible intermediate precursors may be identified using the E. coli cell free extract that contains RNase ( Ikeda 1972 ). Utilizing this RNase, as well as the base and terminus modification systems that have already been defined, the investigation of the sequential order of the maturation of tRNA can be made. The accuracy of these processes can be tested by the biological activity of the tRNA made in vitro.

## References:

- Altamn, S , Nat. New Biol, 229, 19 (1971)
- Bleyman, M. Kondo, M. Hecht, N. and Woese, C., J. Bact 99,  
535 (1969)
- Bernhardt, D., and Darnell, J. E. Jr., J. Mol. Biol. 42, 43 (1969)
- Burdon, R. H., and Clason, A. E., J. Mol. Biol. 39, 113 (1969)
- Cutler, R. G. , and Evans , J. E., J. Mol. Biol. 26, 91 (1967)
- Daniel, V. and Littauer, U. Z., J. Biol. Chem., 238, 2102 (1963)
- Daniel, V. Sarid, S., Beckman, J. S., and Littarer, U. Z.,  
Proc. US Acad. Nat Sci. 66, 1260 (1970)
- Daniel, V., Sarid, S., and Littarer, U. Z., Science, 167, 1682 (1970)
- Fleissner, E. and Borek, E., Proc. US Nat Acad. Sci., 48, 1199 (1962)
- Fleissner, E., and Borek, E., Biochem., 2, 1093, (1963)
- Furth, J. J., Hurwitz, J., Krug, R., and Alexander, M., J. Biol  
Chem., 236, 3317 (1961)
- Hayward, R. S., and Weiss, S. B., Proc. US Nat. Acad. Sci., 55,  
1161 (1966)
- Ikeda, H., Nat. New Biol, 234, 198 (1972)
- Mowsgowitz, D. B., J. Mol. Biol. 50, 143 (1970)
- Pace, B., Peterson, R. L., and Pace, N. R., Proc. US Nat Acad. Sci.  
65, 1097 (1970)
- Russell, R. L., Abelson, J. N., Landy, A., Fetter, M. L., Brenner  
S., and Smith, J. D., J. Mol. Biol. 47, 1 (1970)
- Preiss, J., Dieckman, M., and Berg, P., J. Biol. Chem. 236,  
1748 (1961)
- Smillie, E. J., and Burdon, R. H., Biochem. Biophy. Acta 213,

248 (1970)

Smith, I., Dabnau, D., Morrell, P. and Maumur, J., J. Mol. Biol.

33, 123 ( 1968)

Squires, C., , Conrak, B. Kirschbaum, J., and Carbon, J., Proc.

U S Nat. Acad. Sci. In Press

Weiss, S. B., and Legault-Demare, J., Science, 149, 429 (1965)

Wilson, J. H., Kimm J. S., and Abelson, J. N., ( 1972 ) J. Mol . Biol.

in press.

Wu, M., and Davidson, N. (1973) Manuscript in preparation .

Wu, M., Davidson, N, and Carbon, J. (1972) in press

## PROPOSITION IV

Interactions Among Nerve Growth Factor, Somatomedin, Insulin  
And Insulin Receptor

Hormones have specific interactions with their respective receptors on the target cell membranes. Somatomedin (S) stimulates similar responses in cells as insulin (I) does. S is reported to compete with I for I receptors on cell or membrane fraction. Nerve growth factor (NGF) shows striking structural similarity to proinsulin. This proposition concerns the possible interactions among S, I, NGF, and isolated I receptors and tests the hypothesis that functional similarity reflects structural homology and vice versa.

Hormones bind to special structures, called receptors, on the membranes of their target cells. The interaction between a hormone and its receptor is the initial step in a chain of events which eventually leads to a change in the physiological state of the target cell ( Pastan, Roth, and Macchia 1966 ); for instance, thyroid-stimulating hormone causes an increase in glucose oxidation ( Pastan et. al., 1966). Formation of receptor-hormone complex is reversible and does not involve covalent bonds. The binding has been demonstrated with either whole cells or membrane fractions and radioactive labelled hormones such as ACTH ( Lefkowitz, Roth, Pricer and Pastan 1970 ), insulin (Cuatrecasas, 1971; Freychet, Roth, and Neville, 1971; Vaun Bennett, Cuatrecasas, 1972; Gavin, Roth, Jen and Freychet, 1972 ) and glucagon ( Desbuquois, Cuatrecasas, 1972 ) . The binding is specific in the



sense that only the derivatives of a given hormone or the same hormone from a different source can compete with the radioactive labelled hormone for the binding, whereas other hormones do not ( Lefkowitz et al., 1970; Cuatrecasas, 1971; Freychet et al., 1971; Vaun Bennett, et al., 1972; Gavin, et al., 1972 ). The competition is proportional to the biological activity of the competitor.

The receptors most thoroughly studied are those of insulin. They have been demonstrated in liver cells (Cuatrecasas, 1972 ), adipocytes (Cuatrecasas 1971 ), chondrocytes ( Cuatrecasas, 1972 ) human circulation cells ( lymphocytes, granulocytes, and erythrocytes ) ( Gavin, et al., 1972 ) and fibroblasts with biologically active  $I^{125}$ -insulin. The binding does not inactivate either insulin or its receptor. Recently, Cuatrecasas (1972) has isolated insulin receptors from fat cells and liver cell membrane using a non-ionic detergent Triton X-100. The receptor is a glycoprotein of molecular weight about 300,000 daltons. The binding activity of the receptor is not affected by phospholipase A or C, but is completely destroyed by mild trypsin digestion. Insulin is composed of two peptide subunit held together by an S-S bond, and has a total molecular weight of 57,000 daltons. It stimulates many metabolic processes including uridine uptake, RNA synthesis, polysome formation, protein synthesis, glucose degradation and lipogenesis. Similar metabolic processes are stimulated in nerve cells by nerve growth factor (NGF) ( Angeletti, Gandini-Attardi, Toschi, and Leve-Mantalcini, 1965; Liuzzi, and D'Agnolo, 1969) as well as by insulin at high con-

centration. The similarity between the two hormones does not end here. Recently, the amino acid sequence of the NGF ( composed of two identical subunits, each with 118 amino acid residues ) has been elucidated, and shows great amino acid and structural homology to proinsulin, the precursor ( Angeletti, and Bradshaw, 1971; Frazier, Angeletti, and Bradshaw, 1972 ) of insulin. The hormones and the glands that produce them ( submaxillary and pancreas respectively ) may have a common evolutionary origin ( Frazier, et al., 1972 ).

Insulin-like effects have also been stimulated by another peptide, somatomedin ( Hintz, Clemmons, Underwood, and Van Wyk, 1972 ) of molecular weight about 7,000 daltons. It is formally known as the sulfation factor or the thymidine factor found in serum. Somatomedin has a very broad action spectrum. It stimulates, among other thing, the uptake of  $^{35}\text{SO}_4$  by cartilage, the incorporation of thymidine into DNA, and of uridine into RNA, and many of the same metabolic responses as insulin. Somatomedin can displace  $\text{I}^{125}$ -insulin from its receptors on adipocytes, chondrocytes, and liver membrane ( Hintz, et al., 1972 ). The competition curves with rat adipocytes and rat liver membrane, as measured by the amount of  $\text{I}^{125}$ -insulin bound after addition of the competitor, are parallel to those produced by unlabelled insulin. However, the efficiency of somatomedin is only 2% of that of unlabelled insulin in the case of rat adipocytes, and 0.5% of unlabelled insulin when rat liver membrane is used as a source of receptors. The curves for fetal chicken chondrocytes are not parallel and the efficiency is higher. From these results, somatomedin is believed to compete with insulin for the insulin

receptors. If somatomedin does displace insulin by competing for the insulin receptor, then I predict:

- (A) The receptors for insulin are not identical among the different types of cells of the same animal, as demonstrated by the different efficiencies with which somatomedin displaces insulin from its receptors.
- (B) The functional similarity of insulin, somatomedin, and nerve growth factor reflects their structural homology. In addition, their receptors may be somewhat similar.

However, there is an alternative interpretation for the displacement of insulin by somatomedin: Insulin and somatomedin have separate receptors on the same target cells; the distribution of these receptors is such that the binding of one hormone distort the configuration of the receptor for the other hormone, and increases the  $K_m$  of binding. Since the binding is reversible, large doses of one hormone destabilizes the receptor-hormone complex of the other type and favors the dissociation of the complex. However, the configurational relation of the two kinds of receptor is not necessarily the same on different types of cells, consistent with the different competition curves observed.

The same argument holds for the existence of a divalent receptor with different active sites for insulin and somatomedin. In this case the arrangements of the active sites must be different on the receptors on different cell types. It may even be that one active site ( that for insulin binding ) can be inactivated by mild proteolytic digestion without affecting the other.

To test the various hypotheses, I propose to perform the following experiments :

- (1) In order to determine whether or not the glycoprotein insulin receptors are identical among different cells of the same animal, the receptors will be isolated using phospholipases and TritonX-100 according to the method of Cuatrecasas. Their binding activities towards various insulins and insulin derivatives will be tested. If the receptors show different affinities for the hormones, then they must be different.
- (2) The interaction between somatomedin and insulin receptors will be tested by the ability of somatomedin to inhibit or displace  $I^{125}$ -insulin binding. Binding of somatomedin is thus tested indirectly because pure or radioactive labeled somatomedin is not available. In addition, a difference between receptor and receptor-hormone complex in gel filtration, gel electrophoresis, or isoelectrofocussing patterns may be detected. A change of " behavior " of the isolated insulin receptor in these analyses in the presence of somatomedin then indicates interactions between them. When a positive result is obtained, the possibility of a divalent receptor can next be tested. The isolated insulin receptor is treated with trypsin under mild conditions which completely abolishes its insulin binding ability. The treated insulin receptors will then be analysed by the above mentioned methods in the absence and presence of somatomedin. A change in the results confirms the divalent receptor hypothesis.
- (3) The biological activity of somatomedin is tested on nerve cells. If prediction (B) is correct, somatomedin should stimulate effects

similar to those by nerve growth factor, as insulin does at high concentration.

(4) Similarly, the interaction between nerve growth factor and the isolated insulin receptor will be examined. Due to its larger size, intact nerve growth factor may not bind to the receptors. However, limited proteolytic digestion of the nerve growth factor may produce fragments which will bind to the insulin receptors. The biological activity of such fragments in insulin bioassay systems will subsequently be tested if binding is observed.

receptors. The biological activity of such fragments in insulin bioassay systems will subsequently be tested if binding is observed.

#### References:

Angeletti, R.H., and Bradshaw, R.A., Proc. US Nat. Acad. Sci., 68, 2417 (1971)

Angeletti, P. U.; Gandini-Attardi, D.; Toschi, G., Salvi, M.L., and Levi-Mantalcini, R., Biochim. Biophys. Acta., 95, 111, ( 1965 )

Cuatrecasas, P., Proc. US Nat. Acad. Sci., 68, 1264 (1971)

Cuatrecasas, P., Proc. US Nat. Acad. Sci., 69, 318 (1972)

Desbuquois, B., and Cuatrecasas, P., Nat. New Biol., 237, 202 ( 1972 )

Foppen, F., Liuzzi, A., and D'Agnolo, G., Biochim. Biophys. Acta., 187, 414 ( 1969 )

Frazier, N.A., Angelette, P.H., and Bradshaw, R.A., Science, 176, 482, (1972)

Freychet, P., Roth, J., and Neville, D.M. Jr., Proc. US Nat. Acad. Sci., 68, 1833 (1971)

Hintz, R.L., Clemmons, D.R., Underwood, L.E., and Van Wyk, J.L., Proc. US Nat. Acad. Sci., 69, 2315 (1972)

Gavin, J.R. III, Roth, J., Jen, P., and Freychet, P., Proc. US Nat. Acad. Sci., 69, 747 (1972)

Lefkowitz, R.J., Roth, J., Pricer, W., and Pastan, I., Proc. US Nat. Acad. Sci., 65, 745 ( 1970 )

Pastan, I., Roth, J., and Macchia, V., Proc. US Nat. Acad. Sci., 56,  
1802 ( 1966 )

Vaun Bennett, G., Cuatrecasas, P., Science, 176, 806 ( 1972 )

## PROPOSITION V

ESTROGEN BINDING PROTEINS AND ESTROGEN BINDING TO  
CHROMOSOMES

Multiple forms of estrogen binding proteins have been isolated from calf and rat uteri. Experiments are proposed to study these proteins and to locate the binding sites on the chromosomes.

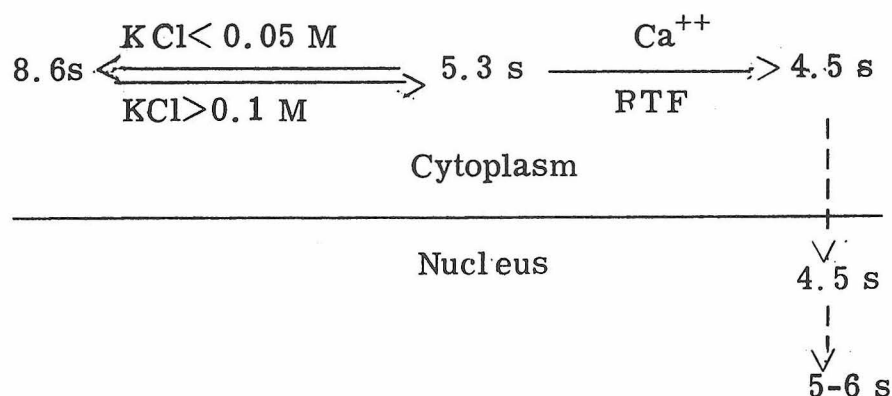
Steroid hormones, estrogens, progestagens, and androgens enter into their respective target ( uterus and prostate ) cells and induce specific responses (1-7). In vivo and in vitro studies indicate that the mechanisms of induction by these hormones are very similar ( 8 ). Using radioactive labeled hormones, multiple macromolecular receptors for the hormones have been isolated from the cytoplasm and the nuclei of the target tissue cells in sucrose gradient centrifugation (2, 3, 6, 8-15 ). Of the three categories of hormones, the interaction between estrogens ( E ) and estrogen binding proteins ( EBP ) from uteri is the most studied and best understood.

The details of the series of events from E uptake to E-EBP formation in the nuclei has been summed up by Puca and associates as follows (16 ):

The E molecule is taken up specifically by its target tissue cell and enters in the cell. Under physiological conditions



(  $\text{KCl} \approx 0.16 \text{ M}$  ), E forms a complex with an EBP in the cytoplasm without covalent bond formation or energy consumption. The complex has a sedimentation coefficient of 5.3 s (molecular weight, 118,000 daltons). It is then converted irreversibly to 4.5 s species (molecular weight, 60,000 daltons) by a  $\text{Ca}^{++}$ -activated protein (called the receptor transforming factor (RTF) of molecular weight at least 100,000 daltons) in a temperature dependent process. Large numbers of the smaller E-EBP complexes enter in the nucleus (2, 17), interact with the chromosome(s), and trigger the syntheses of new species of RNA's and proteins. Estrogens have been shown to bind to isolated chromatin via some acidic, non-histone protein(s) (8, 17-19). EBP's of 4.5 s and 5-6 s have been extracted from the nuclei (13, 16, 20); the former is identical to the cytoplasmic 4.5 s EBP in sucrose gradient centrifugation and gel filtration (16); the latter is presumably the new complex formed between the nuclear 4.5 s EBP and the nuclear acidic protein(s). The 8.6 s EBP first extracted from the cytoplasm turns out to be a dimer of the 5.3 s species formed at a KCl concentration below 0.05 M (15, 16). The relationships among the various EBP's are summarized below:



In sucrose gradient centrifugation and in gel filtration, the cytoplasmic ( 8.6 s ) 5.3 s, 4.5 s and the nuclear 4.5 s EBP's each gives one homogeneous peak. When analysed by isoelectrofocussing, the 8.6 s and the 5.3 s species each gives one peak, but the cytoplasmic 4.5 s EBP is focussed at pH 6.6 and 6.8 ( 15, 16 ). The nuclear 4.5 s species has not been analysed. This result can be interpreted as either: ( 1 ) the RTF is a proteolytic enzyme that splits the 5.3 s chain into two fragments of very similar size. Biologically active 4.5 s EBP can be generated from the 5.3 s EBP by trypsin treatment. However, the trypsin EBP shows multiple peaks in isoelectrofocussing. Tests on the RTF also indicate that RTF is not a trypsin. or ( 2 ) the 5.3 s EBP is made of two different non-covalently linked subunits; the RTF unfolds the subunits and then refolds them into new configurations. In the new configurations, the subunits do not reassociate. Presently it is not known which of the interpretations is correct, nor the number of binding site on each of the EBP's, nor the number and location of the interaction site on the chromosomes. This proposition is directed towards studying these problems.

Experiments proposed are the following:

( 1 ) To obtain highly purified 5.3 s EBP.

Affinity chromatography (21) with estrogen covalently linked to sepharose ( 22 ) will be used. This technique involves covalently linking component 1 of a reversible interaction system to an insoluble medium; component 2 can then be purified in one step by virtue of its specific affinity for the immobilized component

1. The non-covalently adsorbed component 2 can be eluted out by changing either the ionic strength or the pH of the buffer, or by using denaturing agents such as urea or guanidine-HCl. The component 1-2 complex can be eluted by cleaving the covalent bond between the medium and component 1. Sepharose is the best insoluble supporting medium, and has been used successfully to purify various biologically active enzymes ( 23-25 ), hormone ( 26 ), and antibody (27). Polyacrylamide derivatives have also been used ( 28 ). Cuatrecasas has prepared a number of sepharose derivatives capable of reacting with compounds ( proteins ) with alkyl halides, primary aliphatic or aromatic amino groups, or with phenolic, histidyl, imidazole compounds ( 22 ). 3-O-succinyl estradiol has been coupled to a sepharose derivative and used to purify EBP's of calf uterus ( 22 ). The binding to estradiol ( an estrogen ) is so strong that it is impossible to elute the EBP's without denaturing them. Therefore mild base or neutral hydroxylamine is used to cleave the estradiol-sepharose bond to give the E-EBP complex. The presence of the hormone in the EBP will not interfere with some of the experiments I propose to do. When purified, uncomplexed EBP is required, a sepharose-estriol or sepharose-esterone column will be used. Both esterone and estriol are chemically very similar to estradiol ( hence one should be able to covalently link them to sepharose using the same chemistry ) with only 1/10 and 1/ 24 the affinity for EBP's as estradiol ( 15 ). Therefore the EBP can be eluted from the column free of E with either low ionic strength buffer or urea at low concentration ( which does not inactivate the EBP ).

( 2 ) To determine the subunit structure of the purified 5.3 s EBP.

Acrylamide-SDS gel electrophoresis of the above obtained EBP ( complexed or uncomplexed ) will be used. From the number of protein bands and the mobility, one can determine whether the 5.3 s protein is a single chain or it is made of two different non-covalently linked subunits ( of isoelectric points of 6.6 and 6.8 respectively ).

( 3 ) To determine the active fragment or subunit in the 4.5 s EBP.

Affinity chromatography is again used. The 4.5 s EBP converted from the purified, uncomplexed 5.3 s species by the RTF is passed through a sepharose-estradiol ( or a sepharose- estriol ) column after the RTF is removed by gel filtration. Only the active fragment or subunit will be retained. From the amount of protein unbound, one can determine whether only one of the 4.5 s species is biologically active. Another method for solving this problem is to extract the nuclear 4.5 s EBP and compare its behavior in isoelectrofocussing analysis to that of the cytoplasmic 4.5 s EBP. If only one of the 4.5 s proteins is active, we will observe only one peak, either at pH 6.6 or at pH 6.8, since uncomplexed EBP does not enter into the nucleus. As a matter of fact, this is a means to check whether the cytoplasmic and nuclear 4.5 s EBP's are identical as claimed. If no chemical modification has occurred upon entering into the nucleus, the nuclear 4.5 EBP will not have isoelectric point other than pH 6.6 and 6.8.

( 4 ) To determine the number of binding sites on the EBP.

The complex formed between excess  $H^3$ -estradiol and the

above purified, uncomplexed 5.3 s or active 4.5 s EBP is isolated from a sucrose gradient or by gel filtration. The ratio of the amount of estradiol and protein molecules in the complex gives the number of binding site per EBP.

( 5 ) To isolate the acidic, non-histone binding protein(s) in the nudeus,

Affinity chromatography involving three components, sepharose-estradiol, 4.5 s EBP, and extracted nuclear acidic proteins, added in that order will be used. The bound acidic protein(s) will then be eluted and further characterized.

( 6 ) To locate the binding site(s) in the nucleus.

To determine the number and location of interaction site on the chromosome, autoradiography will be used. This technique has been used to show that estradiol is taken up by its target tissue and is mostly localized in the nucleus ( 29 ). However no deeper investigation has been made in spite of the fact that a superb no fixing - no embedding- freeze dry- dry mount autoradiography technique has been developed and demonstrated on  $H^3$ -estradiol-uterus tissue sections ( 29, 30 ). Maio and Schildkraut have isolated metaphase chromosomes from purified nuclei of several culture cell lines in a neutral hypotonic buffer ( 31, 32 ). This neutral buffer will be used to prepare metaphase nuclei and chromosomes from rat uteri. A chromosome squash (33) will be made from the purified nuclei on a glass slide in the neutral buffer. It is freeze-dried to fix ( to some extent ) the chromosomes onto the slide. Alcohol, acid, or formaldehyde fixation should be avoided as it probably will destroy the binding activity of the chromosomes. A highly radioactive  $H^3$ -estradiol-4.5 s

EBP complex will then be added and incubated for complex formation. The excess unbound E-EBP is washed off. Formaldehyde or glutaldehyde fixation may be performed at this point. The sample is then prepared for autoradiography. After exposure and developing, it is stained and examined with a light microscope. From the distribution of the silver grains, one can locate the binding site(s) on the chromosome(s). Alternatively, electron-dense spheres of proper size ( 300-1000Å ), such as the NH<sub>2</sub>- derivative of the polyacrylate beads made in this laboratory by Drs. N.D. Hershey and T.R. Broker can be coupled to the 3-O-succinyl-estradiol in the presence of a carbodiimide. The sphere-estradiol-EBP complex will then be used to bind to the chromosome squash. Electron microscope grids can be made after the unbound complexes are washed off. From the location of the sphere which appears opaque in the electron microscope, one can locate the binding site. The binding experiment can also be performed with purified metaphase chromosomes prepared from the neutral buffer. The binding will be carried out in solution. The chromosomes can be pelleted down and separated from the unbound E-EBP with or without the spheres attached. The complex will be prepared for autoradiography or direct electron microscopy with or without fixation. I also propose to arrest the uterus cells in metaphase with vinblastine or colchicine after H<sup>3</sup>- estradiol is administered either in vivo ( with life animal ) or in vitro ( with excised uterus ). The isolated nuclei will be fixed with formaldehyde or glutaldehyde before making the chromosome squash. Autoradiography will then be performed on these chromosomes.

## References :

- 1 Jensen, E.V., and Jacobson, H.I. (1962) Recent Prog. Hormone Res., 18, 387
- 2 Gorski, J., Toft, D., Shyamala, G., Smith, D., and Notides, A. (1968) Recent Prog. Hormone Res., 24, 45
- 3 Fang, S., and Liao, S. (1971) J. Biol. Chem., 246, 16
- 4 Swaneck, G.E., Chu, L.L.H., and Edelman, J.S. (1970) J. Biol. Chem., 24, 5382
- 5 Mainwaring, W.I.P., (1969) J. Endocrinol., 44, 323
- 6 O'Malley, B.W., and Toft, D., (1971) J. Biol. Chem., 246, 1117
- 7 Steggles, A.W., Spelsberg, T.C., and O'Malley, B.W. (1971) Biochem. Biophys. Res. Comm., 43, 20
- 8 Steggles, A.W., Spelsberg, T.C., Glasser, S.R., and O'Malley B.W., (1971) Proc. US Nat. Acad. Sci., 68, 1479
- 9 Toft, D., and Gorski, J. (1966) Proc. US Nat. Acad. Sci., 55, 1574
- 10 Wotiz, H.S., and Wotiz, H.H. (1967) Steroids, 10, 635
- 11 Jensen, E.V., Hurst, D.J., DeSombre, E.R., and Jungblut, P.W. (1967) Science, 158, 385
- 12 Jensen, E.V., Suzuki, T., Kawashima, T., Stumpe, W.E., Jungblut, P.W., and DeSombre, E.R. (1968) Proc. US Nat. Acad. Sci., 59, 632
- 13 Jensen, E.V., Suzuki, T., Numato, M., Smith, S., and DeSombre, E.R. (1966) Steroids, 13, 417
- 14 Musliner, T.A., Chader, G.J., and Villet, C., (1970) Biochem. 9, 4448

- 15 Puca, G.A., Nola, E., Sica, V., and Bresciani, F. (1971)  
Biochem. 10, 3769
- 16 Puca, G.A., Nola, E., Sica, V., and Bresciani, F. (1972)  
Biochem. 11, 4157
- 17 Noteboom, W.D., and Gorski, J., (1965) Arch. Biochem.  
Biophys., 111, 559
- 18 King, R.J. B., and Inman, D.R. (1966) J. Endocrinol. 35, xxvi
- 19 Shyamala, G., and Gorski, J., (1967) J. Cell Biol., 35, 125A
- 20 Puca, G.A., and Breciani, F., (1969) Nature, 223, 745
- 21 Cuatrecasas, P. (1971) Ann. Rev. Biochem. 40, 753
- 22 Cuatrecasas, P. (1970) J. Biol. Chem., 245, 3059
- 23 Porath, J., Axén, R., and Ernback, S. (1967) Nature, 215, 1491
- 24 Cuatrecasas, P., Wilchek, M., and Anfinsen, C. B., (1968)  
Proc. Nat. Acad. Sci., 61, 636
- 25 Chan, W.W.C., and Takahashi, M. (1969) Biochem. Biophys.  
Res. Comm., 37, 272
- 26 Gospodarowicz, D. (1972) Biol. Chem. 247, 649
- 27 Cuatrecasas, P., (1969) Biochem. Biophys. Res. Comm., 35, 531
- 28 Inman, J. K., and Dintzis, H. M. (1969) Biochem. 8, 4074
- 29 Stumpf, W. E., (1968) Endocrinol., 83, 777
- 30 Stumpf, W. E., and Roth, L. J. (1966) J. Histochem. Cytochem.  
14, 274
- 31 Maio, J. J., and Schildkraut, C. L. (1967) J. Mol. Biol. 24, 29
- 32 Maio, J. J., and Schildkraut, C. L. (1969) J. Mol. Biol., 40, 203
- 33 John, H. A., Birnstiel, M. L., and Jones, K. W., (1969) Nature,  
223, 582



## PROPOSITION VI

SYNTHESES OF DNA COMPLEMENTS OF NATURAL RNA's –  
A GENERAL APPROACH\*

Syntheses of DNA complements of natural RNA's are proposed to be carried out by using the reverse transcriptases from RNA viruses in the presence of T<sub>4</sub> gene 32 protein which melts out secondary structures in single-stranded DNA and RNA.

Reverse transcriptases ( RNA-dependent DNA polymerases ) have been isolated from virions of a number of RNA viruses ( 1,2 ) The enzyme uses a number of viral RNA's, E. coli tRNA's and synthetic homopolymers as template in vitro ( 3-9). It does not initiate new DNA chains, but incorporates deoxynucleotides to the 3'-OH ends of the primer strands. It can also use double-stranded DNA with template-primer structure as template-substrate (10). Short single-stranded regions in the double-stranded DNA are completely filled in. However double-stranded DNA with extensive single-stranded gaps is not a good template. Single-stranded and completely double-stranded DNA's with or without nicks are not used by the enzyme. The products made on an RNA template are DNA-RNA hybrid, single- and double-stranded DNA. When DNA is used as a template substrate, the product is duplex DNA. Apparently, the reverse transcriptase has both RNA-dependent and DNA-dependent DNA polymerase activities; it uses RNA as an initial template with RNA or DNA as primer to make a DNA-RNA

hybrid. The initial product is then served as template for the synthesis of double-stranded DNA (5) .

It is speculated that in vivo the reverse transcriptase makes DNA copies from the viral RNA; the DNA complement is then integrated into the host chromosome. However, in all in vitro studies, the DNA products made on either endogeneous or exogeneous RNA templates are short ( 6-8 s ) even though the RNA templates used are rather long ( 70 s ).

DNA complements ( ~ 8 s ) of duck and rabbit hemoglobin mRNA's ( 10 s ), ( 11 ) ; human globin mRNA ( 10 s ) (12), and vaccinia mRNA ( 'made in vitro, ~ 10-12 s ) ( 13 ) have been made in several laboratories, using the reverse transcriptases from different RNA virions in the presence of oligo dT. Oligo dT is used as primers in these studies based on the observation that most mRNA's of eucaryotes have a stretch (50-200 bases) of poly A at the 3' ends ( 14-16 ).

The above observations indicate: (1) The presence of a primer is obligatory for the activity of the reverse transcriptase. (2) The secondary structure of the template may play an important role in the enzyme activity. The failure of the reverse transcriptase to synthesize, in vitro, long DNA complement on long endogeneous or exogeneous natural RNA template may be due to its inability to overcome the intramolecular forces of the template. The fact that the enzyme uses well double-stranded DNA with short, but not long single-stranded regions as template supports this hypothesis. The reasons that RNA's other than 10 s extracted from polyribosomes of duck or rabbit reticulocytes are not used by the enzyme as template (11) may be either due

to the lack of proper primer and / or the secondary structure of the RNA's.

The problem of providing a proper primer can be solved by attaching a stretch of poly A to the 3' end of the RNA using the E.coli polynucleotide phosphorylase or the polyA polymerase present in animal cells as proposed by Zassenhaus and Kates ( 13 ).

To overcome the secondary structure in the template, I propose to use  $T_4$  gene 32 Protein.

$T_4$  gene protein binds to single stranded DNA in vitro (17). The DNA-protein complex assumes an extended appearance. The protein can also bind to double-stranded DNA at the AT rich region, and induces local denaturation. In vivo,  $T_4$  gene 32 protein is involved in  $T_4$  replication (18). It has also been shown that in vitro, the rate of DNA synthesis by  $T_4$  DNA polymerase is stimulated 5 to 10 fold when  $T_4$  gene 32 protein is added (19).

In vitro experiments show that  $T_4$  gene 32 protein also binds to single-stranded RNA (20). The affinity for RNA is much smaller than that for DNA. Nevertheless, the secondary structures of R 17 RNA, and mRNA's of  $T_7$  and SV40 DNA are destroyed when  $T_4$  gene 32 protein is added and fixed by glutaldehyde as observed by electron microscope. The RNA-protein complex also has an extended contour.

In summary,  $T_4$  gene 32 protein (1) can melt out secondary structures in single-stranded DNA or RNA, (2) not only does not interfere with the activity of DNA polymerase, but instead stimulate its activity.

It is possible that  $T_4$  gene 32 protein will not interfere with the

activity of the reverse transcriptase either. If this is true, then T<sub>4</sub> gene 32 protein will definitely stimulate the activity of the reverse transcriptase by melting out the secondary structures of the template.

The experiment proposed is to compare the size of the products made by the reverse transcriptase on long viral RNA's in the absence and in the presence of the T<sub>4</sub> gene 32 protein.

There is great significance if the reverse transcriptase indeed makes long DNA complement of the RNA template in the presence of the T<sub>4</sub> gene 32 protein; that is, we can make ( and amplify ) in vitro " genes " from any mRNA's ( preferably with poly A at the 3' end). Once large quantities of the DNA complements of the mRNA's in hand, we can carry out many experiments; for instance, to establish precursor-product relationship of the RNA's, to detect viral mRNA in cells infected with RNA viruses, and to map the messages carried in the RNA on the chromosomes in either procaryotes or eucaryotes by electron microscopy heteroduplex method (21) or by in situ hybridization to metaphase chromosomes ( or nuclei ) and autoradiography (22, 23). It is also conceivable that one may want to make large quantities of proteins encoded in the amplified DNA complements in in vitro protein synthesis systems for medical uses.

At this point, the task will be to identify the messages encoded in the mRNA that one isolates.

\* The title is borrowed from Spiegelman, et al., (8).

## References:

- 1 Baltimore, D. (1970) *Nature*, 226, 1209.
- 2 Temin, H.M., and Mizutani, S. (1970) *Nature*, 226, 1211.
- 3 Rokutanda, M., Rokutanda, H., Green, M., Fujinaga, K., Ray, R.K., and Gurgo, C., (1970) *Nature*, 227, 1026.
- 4 Spiegelman, S., Burny, A., Das, M.R., Keydar, J., Schlom, J., Travricek, M., and Watson, K. (1970) *Nature*, 227, 563.
- 5 Faras, A., Fanshier, L., Garapin, A-C., Levinson, W., and Bishop, J.M., (1971) *J. Virol.* 7, 539.
- 6 Verma, I.M., Meuth, N.L., Bromfeld, E., Manly, K.F., and Baltimore, D. (1971) *Nature new Biol.* 233, 131.
- 7 Duesberg, P., Helm, K.V.D., and Canaani, E. (1971) *Proc. US Nat. Acad. Sci.*, 68, 747.
- 8 Spiegelman, S., Watson, K.F., and Kacian, D.L., (1971) *Proc. US Nat. Acad. Sci.*, 68, 2843.
- 9 Leis, J.P., and Hurwitz, J. (1972) *J. Virol.* 9, 130.
- 10 Hurwitz, J., and Leis, J.P. (1972) *J. Virol.* 9, 116.
- 11 Verma, I.M., Temple, G.F., Fan, H., and Baltimore, D. (1972) *Nature New Biol.* , 235, 163.
- 12 Kacian, D.L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L., and Marks, P.A. (1972) *Nature New Biol.*, 235, 167.
- 13 Zassenhaus, P., and Kates, J. (1972) *Nature New Biol.* 238, 139.
- 14 Darnell, Jun., J.E., Wall, R., and Tushinski, R.J., (1971) *Proc. US Nat. Acad. Sci.*, 68, 1321.
- 15 Edmonds, M., Vaughan, Jun., M.H., and Nakamoto, H. (1971)

- Proc. US Nat. Acad. Sci., 68, 1336.
- 16 Philipson, L., Wall, R., Glickman, G., and Darnell, J.E. (1971)  
Proc. US Nat. Acad. Sci., 68, 2806.
- 17 Delius H., and Mantell, N.J. (1972) J. Mol. Biol., 67, 341.
- 18 Sinha, N., and Snustad, D.P., (1971) J. Mol. Biol. 62, 267.
- 19 Huberman, J., Kornberg, A., and Alberts, B. (1971) J. Mol. Biol.  
62, 39.
- 20 Delius, H., and Westphal, H. (1972) J. Mol. Biol., in press.
- 21 Davis, R.W., Sinmon, M.N., and Davidson, N. (1971) IN  
Methods in Enzymology, vol. 21D, p. 413.
- 22 Hearst, J.E., and Blotchan, M. (1971) Ann. Rev. Biochem. 39,  
151.
- 23 John, H.A., Birnstiel, M.L., and Jones, K.W., (1969) Nature,  
223, 582.